



Stack Overflow Scalability

โดย

นายอัศวรัตน์ อัมพรณไชย

โครงการในรายวิชาคพ. 447 นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

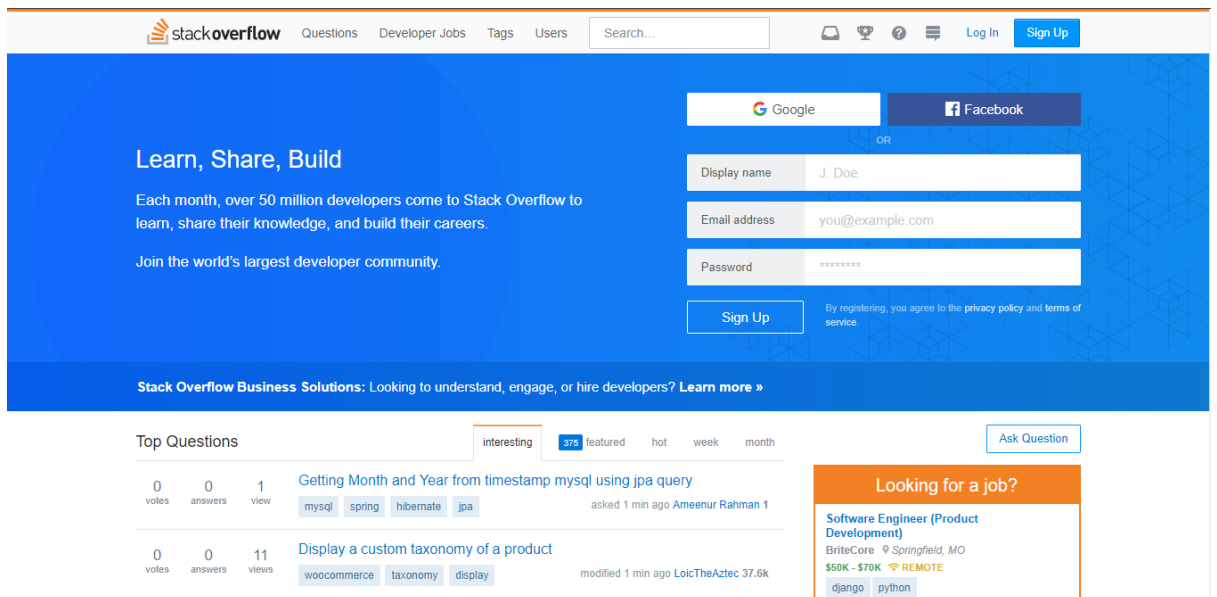
วิทยาศาสตร์บัณฑิตสาขาวิชาวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์

ปีการศึกษา 2560

1. Stack Overflow ให้บริการอะไร

Stack Overflow เป็นบริษัทในเครือ Stack Exchange ที่ให้บริการเกี่ยวกับ Q&A site เพื่อเรียนรู้ และแบ่งปันความรู้โดยคำถามบน Stack Overflow ส่วนมากเป็นคำถามเกี่ยวกับการพัฒนาโปรแกรมคอมพิวเตอร์ หน้าหลักของ Stack Overflow ดังรูปที่ 1



รูป 1 Stack Overflow homepage

ปี 2014 Stack Overflow มีอัตราการชมที่ 560 ล้านครั้งต่อเดือนของผู้ใช้งานจำนวนกว่า 4 ล้านคน จำนวนกระทู้คำถามมีมากกว่า 8 ล้านกระทู้ และมีจำนวนคำตอบมากกว่า 40 ล้านคำตอบ รวมถึงอัตราการเติบโตของเว็บไซต์มากกว่า 100% ต่อปี ทำให้เห็นว่า ถ้าการทำ Scalability ไม่ดีจะไม่สามารถรองรับการใช้งานอย่างมหาศาลได้

2. ปัญหาการ Scalability วิธีการแก้ และผลลัพธ์

การทำ Scalability ครั้งใหญ่ของ Stack Overflow เกิดขึ้น 3 ครั้งในปี 2009 2011 และ 2014 ตามลำดับ รายละเอียดของการเปลี่ยนแปลงแต่ละครั้งมีดังนี้

ปี 2009 Stack Overflow มีผู้ใช้อยู่ประมาณ 3 ล้านคนต่อเดือน และมีการเข้าถึงประมาณ 16 ล้านครั้งต่อเดือนโดย 86% ของปริมาณการใช้งานมาจาก Google 30% ของโปรแกรมเมอร์ทั่วโลกใช้ Stack Overflow Hardware ของ Stack Overflow ที่ใช้มีจำนวนเพียง 4 เครื่องสำคัญคือ Web Server 2 เครื่องใช้ Microsoft IIS Database Server 1 เครื่องใช้ Microsoft SQL Server และ NAS สำหรับ Backup ข้อมูล 1 เครื่อง

การเพิ่ม Scalability ครั้งนี้ Stack Overflow ใช้เทคนิคการ Vertical Scaling ด้วยการเพิ่มความเร็วของแรม Database server เป็น 48 GB เพิ่มความเร็วของหน่วยประมวลผลกลางของ Database server จาก 1.86 GHz เป็น 2.5 GHz และ 3.5 GHz ตามลำดับเพื่อเพิ่มประสิทธิภาพของการ Query ข้อมูลเพื่อแก้ปัญหาคอขวดของ Database ที่ใช้เวลาประมาณ 90% ของการทำงานทั้งหมด และการออกแบบ Database บางส่วนมีต้นแบบมาจาก Wikipedia ส่งผลให้เกิดการใช้ทรัพยากรมากเกินไปจนความจำเต็ม จึงต้อง Refactor ด้วยการหลีกเลี่ยงการ Query ที่มากเกินไปทำให้ลดทรัพยากรที่มากจำเป็นได้ แต่มีบางปัญหาที่ยังไม่ได้แก้คือ ปัญหาการค้นหาแบบ Full Text Search ซึ่งแต่เดิมใช้ Microsoft SQL Server มีปัญหาที่ราคาค่าใช้จ่าย รวมถึงประสิทธิภาพในการทำงานเชิงลึกทำให้ทีมพัฒนาสนใจ Lucene และปัญหาของ Relational SQL มีข้อผิดพลาดมากมาย แต่ก็สามารถใช้งานได้สำหรับงานทั่วไป ถึงแม้ว่า NoSQL จะเป็นทางเลือกหนึ่งแต่ด้วยความยากต่อการบำรุงรักษา และค่าใช้จ่ายที่สูงทำให้ Stack Overflow ยังใช้ SQL อยู่

ผลที่ได้จากการ Scale ในปี 2009 คือ Stack Overflow สามารถรองรับการใช้งาน จำนวนของผู้ใช้ได้มากขึ้น ปัญหาคอขวดของ Database ดีขึ้น

ปี 2011 เป็นปีที่ Stack Overflow ปรับปรุงครั้งใหญ่ที่สุดเพราะระบบเก่าที่เคยใช้ตั้งแต่ปี 2009 ไม่สามารถรองรับจำนวนการใช้งานได้อีก ในปี 2011 Stack Overflow มีจำนวนการเข้าชมถึง 95 ล้านครั้งต่อเดือนหรือคิดเป็นประมาณ 6 เท่าตัวของปี 2009 และจำนวนผู้ใช้งานถึง 16 ล้านคน ส่งผลให้ปริมาณการร้องขอบริการสูงถึง 800 ครั้งต่อวินาที

การเปลี่ยนแปลงในปีนั้นเริ่มจากการเปลี่ยนระบบบางส่วนมาใช้ Linux จากที่เป็น Windows Stack ทั้งหมด การใช้ HAProxy เป็น Load Balancer แทน Windows NLB เดิมเพราะ HAProxy มีราคาที่ถูก และง่ายกว่า เริ่มใช้ Redis ที่เป็น NoSQL บน Caching Layer เพราะ Redis เป็น Distributed memory caching โดยเก็บข้อมูลในรูปแบบ key-value ที่สามารถแบ่งปันข้อมูลระหว่าง Server ได้ง่าย เปลี่ยนระบบค้นหาจากระบบค้นหาของ SQL Server มาเป็น Lucene เนื่องจากการค้นหาแบบ Full Text Search ของ SQL Server ประยุกต์ใช้งานไม่ค่อยได้ มีข้อบกพร่องเยอะ ในปีนี้ Stack Overflow เริ่มทำ Cache ดังนี้

1. Local Cache เก็บข้อมูลประเภท User Session, View Count เพื่อลดปริมาณการค้นหา Cache
2. Site Cache เก็บรหัสประจำตัวของคำถาม คะแนนประเมินการยอมรับของผู้ใช้ใน Redis
3. Global Cache เก็บโควตาของการใช้งาน API และที่ชั้นนี้เก็บข้อมูลใน Redis เช่นเดียวกัน

และที่สำคัญ Stack Overflow เปลี่ยน และเพิ่ม Hardware จำนวนมากคือ

1. Web Server ที่ใช้ Microsoft IIS จำนวน 3 เครื่อง
2. Database Server บน MS SQL Server จำนวน 2 เครื่อง
3. HAProxy Server บน Ubuntu Server จำนวน 2 เครื่อง
4. Redis Server บน CentOS จำนวน 2 เครื่อง
5. Backup Server ด้วย Bacula จำนวน 1 เครื่อง
6. Nagios and Logs Server จำนวน 1 เครื่อง
7. Domain Controllers สำหรับให้ผู้ใช้ใช้ทรัพยากรการข้าม Server จำนวน 2 เครื่อง
8. Routers จำนวน 2 ตัว

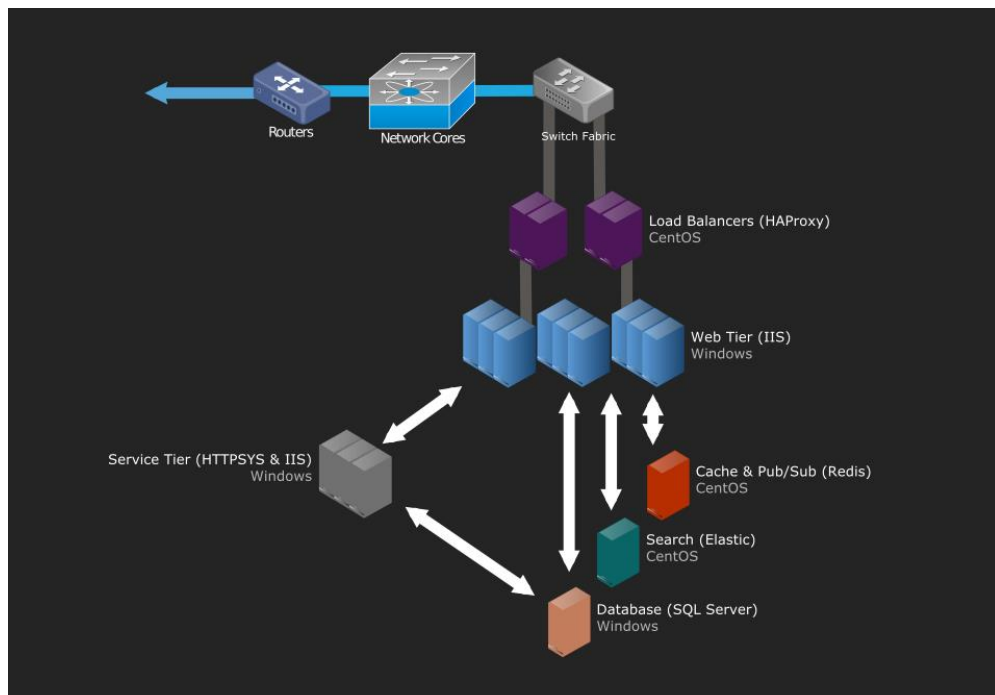
ผลที่ได้จากการ Scale ในปี 2011 คือ Stack Overflow สามารถรองรับการใช้งาน จำนวนของผู้ใช้ได้มากขึ้น การเก็บ Cache ทำให้ Server สามารถทำงานได้เร็วขึ้น ประสิทธิภาพของการทำ Full-text Search ดีขึ้นเพราะ Lucene สนับสนุนการทำ Full-text Search ดีกว่า Microsoft SQL Server การใช้ Load Balancer ทำให้ Server สามารถทำงานได้มีประสิทธิภาพมากขึ้นเพราะการทำงานของ Server จะถูกเฉลี่ยกัน

ปี 2014 การเติบโตของ Stack Overflow โตขึ้นแบบก้าวกระโดดประมาณ 100% ต่อปีโดยมีจำนวนการเข้าชมถึง 560 ล้านครั้งต่อเดือนซึ่งเยอะกว่าปี 2011 เกือบ 6 เท่าตัวกลายเป็นเว็บไซต์ที่ Traffic อันดับที่ 54 จำนวนผู้ใช้มีถึง 26.9 ล้านคนจาก 242 ประเทศ อัตราการใช้งานสูงสุดสร้างคำร้องขอบริการสูงถึง 2,600 – 3,000 ครั้งต่อวินาทีทำให้ Stack Overflow ต้องขยายระบบอีกครั้ง

Stack Overflow เพิ่มจำนวน Server ทั้งหมดเป็น 25 เครื่องประกอบด้วย

1. Web Server ที่ใช้ Microsoft IIS จำนวน 11 เครื่อง
2. Load Balancer ใช้ HAProxy จำนวน 2 เครื่องทำงานเพียง 1 เครื่อง
3. Database Server ใช้ Microsoft SQL จำนวน 4 โหนด
4. Application Server 3 เครื่องเพื่อทำงานเกี่ยวกับ Tag
5. ElasticSearch 3 เครื่องเพื่อการค้นหาข้อมูลใน Stack Overflow
6. Redis จำนวน 2 เครื่องเพื่อจัดการ Cache

ข้อมูล SQL ทั้งหมดเก็บลง SSD ขนาด 2 TB Web Server แต่ละตัวมี SSD ขนาด 320 GB 2 ตัวเก็บข้อมูลแบบ RAID 1 ระบบค้นหา ElasticSearch ใช้ SSD ขนาด 300 GB เพื่อเก็บดัชนีของข้อมูล การเพิ่มจำนวน Server ทำให้ CPU load อยู่ในเกณฑ์ต่ำ Database Server ถูกใช้งานเพียง 10% ยกเว้นตอน Backup ข้อมูล และใช้ Router (Cisco 3945), Networks (Nexus 5596 + Fabric Extenders) และ Firewall (Cisco 5525-X ASAs) อย่างละ 2 ตัว โครงสร้างของสถาปัตยกรรมของ Stack Overflow ดังรูปที่ 2



รูป 2 Stack Overflow infrastructure

Stack Overflow เปลี่ยนการเก็บ Cache เป็น 5 ระดับดังนี้

1. Network Cache เช่น การเก็บ Cache ที่ Browser การใช้ CDN และ Proxy
2. Server Cache โดยใช้การเก็บ Cache ของ .NET (HttpRuntime.Cache) และ In-memory ของ Server
3. การ Cache ที่ Redis ด้วยการเก็บ Cache ในรูปแบบ Key-value ทำให้สามารถแชร์ Cache ระหว่าง Server ได้ง่าย
4. SQL Server Cache ด้วยการเก็บข้อมูลทุกอย่างใน Database ให้เป็น Cache
5. SSD Cache เก็บข้อมูลที่ถูกระบุเรียกใช้ถูก Cache ไว้ และ Cache ระดับนี้ถูกระบุเรียกใช้ต่อเมื่อ Cache ระดับ 4 ยังไม่มีข้อมูล

Stack Overflow เริ่มต้นใช้ Cloud เพื่อการให้บริการที่รวดเร็วเพราะ Cloud ทำให้ Server Render เร็วขึ้นทำให้สามารถช่วยปัญหาคอขวดของ Server ได้ ถึงแม้ว่า Cloud ไม่สามารถควบคุมความเร็วของเครือข่ายให้ดีกว่า 50 ms ได้แต่ก็เป็นทางเลือกที่ดีที่สุดของ Stack Overflow

Stack Overflow ตั้งเป้าหมายด้านประสิทธิภาพในการโหลดหน้าหลักไว้ที่ 50 ms แต่สามารถทำได้ด้วยความเร็ว 28 ms ด้วยเหตุผลหลักที่ว่า Server มีค่า Utilization ที่ต่ำ Web Server มีอัตราการใช้งาน CPU เพียง 10 – 15% และ SQL Server มีอัตราการใช้งาน CPU เพียง 5 – 10% เท่านั้น

แนวทางการพัฒนาของ Stack Overflow คือ Move fast and break things เริ่มด้วยการพัฒนาความสามารถใหม่เข้ามาใน Stack Overflow หลังจากนั้นทีมพัฒนาพัฒนา และทดสอบบนเครื่องตัวเองโดยการทดสอบของ Stack Overflow ไม่ค่อยเยอะเพราะ Code ส่วนใหญ่เป็น Static Code ถ้าผ่านเกณฑ์การทดสอบผ่าน ทีมพัฒนาจะ Push Code ไปยัง Repository TeamCity ที่เป็นระบบ CI (Continuous Integration) ที่ติดตั้งไว้ใน Stack Overflow จะ Build, Test และ Deploy เมื่อผ่านทุกขั้นตอนแต่ตอนใช้งานจริง ระบบเกิดผิดพลาด ทีมพัฒนาย้อนเวอร์ชันของ Code ไปยัง Code เวอร์ชันก่อนหน้าแล้วแก้ไข Code เวอร์ชันล่าสุด

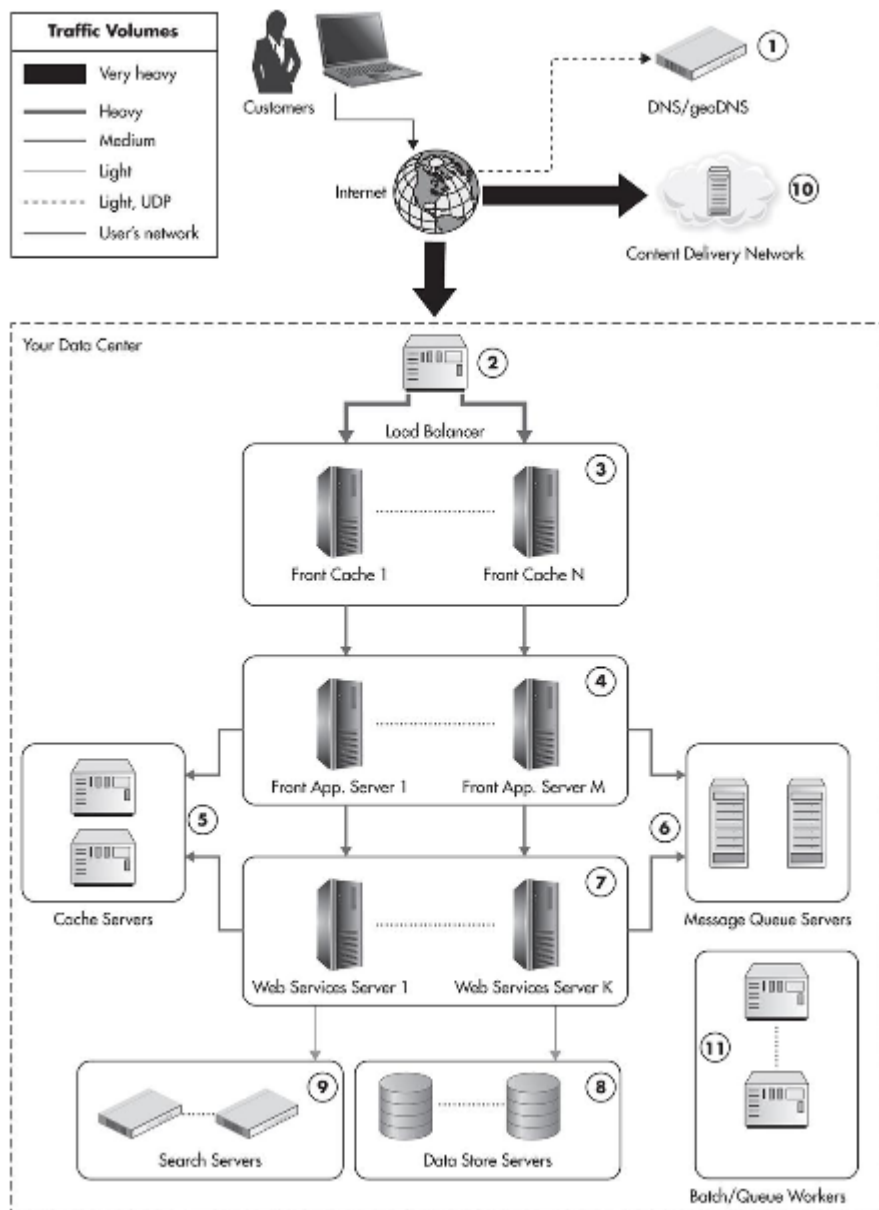
ผลที่ได้จากการ Scale ในปี 2014 คือ Stack Overflow สามารถรองรับการใช้งาน จำนวนของผู้ใช้ได้มากขึ้น ประสิทธิภาพการทำงานโดยรวมดีขึ้นมาก การใช้ Cloud ทำให้ Stack Overflow สามารถบริการผู้ใช้ได้ดีขึ้นเพราะลดปัญหาคอขวดของ Server

3. สรุปความสัมพันธ์กับเนื้อหาที่เรียนในชั่วโมง

จะเห็นว่า StackOverflow ทำ Scale ทั้ง 2 แบบคือ Vertical Scaling และ Horizontal Scaling โดย Vertical Scaling ใช้ตอนเพิ่มที่เก็บข้อมูลที่ทำเป็น RAID 1 ในปี 2011 เพิ่มประสิทธิภาพของ I/O ด้วยการเปลี่ยนไปใช้ SSD ลดระยะเวลาทำงานของ I/O ด้วยการเพิ่ม RAM ของ Server ในปี 2009 และการเพิ่มความเร็วของหน่วยประมวลผลกลาง และแบบ Horizontal Scaling ใช้ตอนเพิ่มจำนวน Server ในปี 2011 และปี 2014

รูปที่ 3 แสดงภาพรวมของระบบที่ Scale แล้ว เมื่อนำมาเชื่อมโยงกับโครงสร้างพื้นฐานของ Stack Overflow จะเห็นว่า โครงสร้างพื้นฐานของ Stack Overflow เหมือนกับรูปนี้เกือบทั้งหมด ดังนี้

1. DNS ของ Stack Overflow ไม่ได้ใช้ GeoDNS เนื่องจาก Server ตั้งอยู่ที่อเมริกา
2. Load Balancer ของ Stack Overflow ใช้ HAProxy Server โดยเริ่มใช้ตอนทำ Scale ในปี 2011
3. Front Cache ของ Stack Overflow ใช้ .NET และ In-memory ของ Server
4. Front Application Server ของ Stack Overflow ใช้ Microsoft IIS Server โดยเขียนด้วยภาษา ASP.Net
5. Cache Server ของ Stack Overflow ใช้ Redis เนื่องจากการเก็บข้อมูลของ Redis เก็บในรูปแบบ Key-value ทำให้สามารถแชร์ข้อมูลระหว่าง Server ได้ง่ายขึ้น
6. Manage Queue Server ของ Stack Overflow ไม่ปรากฏในเอกสารอ้างอิง
7. Web Services Server ของ Stack Overflow ใช้ Microsoft IIS Server เช่นเดียวกับ Front Application Server แต่เพิ่ม Http.sys เพื่อให้สามารถให้บริการ API ได้
8. Data Store Server ของ Stack Overflow ใช้ Microsoft SQL Server ตั้งแต่ปี 2009
9. Search Server ของ Stack Overflow เปลี่ยนมาตั้งแต่ Microsoft SQL Server มาใช้ Lucene และในปี 2014 เปลี่ยนมาใช้ Elasticsearch ที่ใช้พื้นฐานมาจาก Lucene
10. CDN ของ Stack Overflow เริ่มใช้ในปี 2014 เพื่อให้ผู้ใช้สามารถหาคำตอบของคำถามได้เร็วขึ้น
11. Batch/Queue Workers ของ Stack Overflow ไม่ปรากฏในเอกสารอ้างอิง



រូប 3 A Data Center Infrastructure

4. Reference

1. <http://highscalability.com/blog/2009/8/5/stack-overflow-architecture.html>
2. <http://highscalability.com/blog/2011/3/3/stack-overflow-architecture-update-now-at-95-million-page-vi.html>
3. <http://highscalability.com/blog/2014/7/21/stackoverflow-update-560m-pageviews-a-month-25-servers-and-i.html>
4. <http://www.somkiat.cc/development-of-stackoverflow/>
5. <https://www.infoq.com/presentations/stack-exchange>
6. <https://nickcraver.com/blog/2016/02/17/stack-overflow-the-architecture-2016-edition/>
7. <https://stackoverflow.blog/2014/02/19/2013-stack-overflow-user-survey-results/>