

OpenSource

THE COMPLETE MAGAZINE ON OPEN SOURCE

For You

Volume: 04 | Issue: 02 | Pages: 108 | November 2015

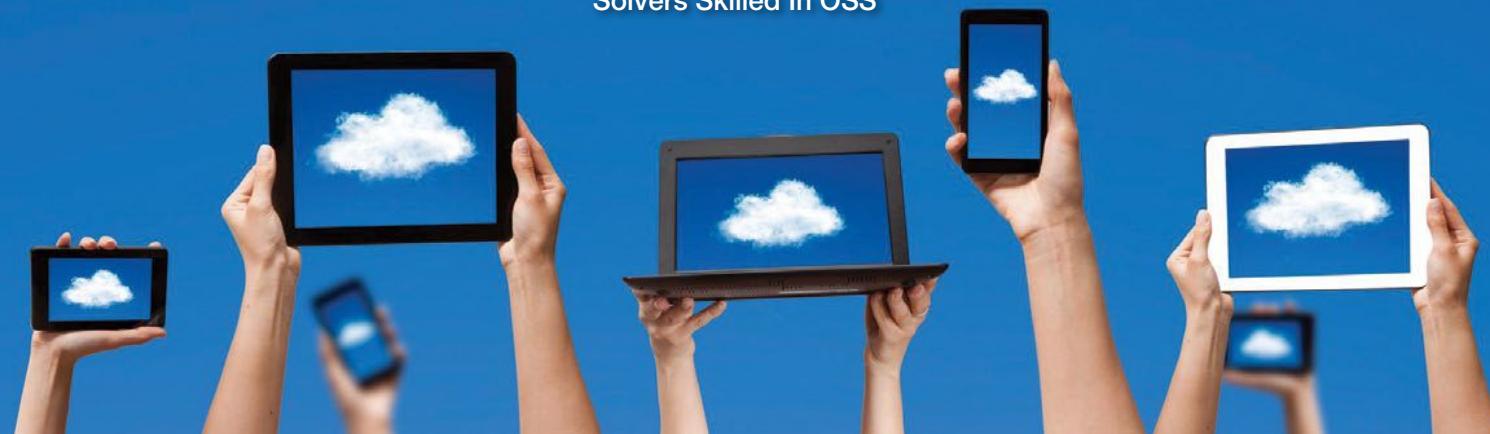
An **EFY GROUP** Publication

CLOUD MANAGEMENT MADE EASY

- Puppet: Your Configuration Manager
- Shorter Software Delivery Timelines With Cloud-IDE
- Scalr: The Efficient Cloud Management Tool

An Interview With Sunil Gupta,
Executive Director And President, Netmagic

Zopper: Looking For Problem
Solvers Skilled In OSS



The Community Is Meeting The Industry!

Have You Marked The Dates On Your Calendar?

12th Edition

OPEN
SOURCE INDIA

NIMHANS Convention Center

19 - 20
November
2015

BENGALURU

Asia's largest conference on Open Source

Change thinking, Increase profitability

Let's THINK BEYOND A BOX



Microsoft Certified Hosting Provider



SAP Certified Cloud & Data Center



ISO 9001 : 2008,



ISO 27001 : 2005,



ISO 20000 - 1 : 2005,



Rating : SE BB



Mumbai | Delhi | Bengaluru | Nashik | Leeds (UK)

ESDS SOFTWARE SOLUTION PVT. LTD.

Toll Free No. : 1800 209 3006 | Fax : +91 95 95 247 247
Email : relationship@esds.co.in | Website : www.esds.co.in

India's first
mega event for all
the stake-holders of
**Internet
of Things**
ecosystem to
come together.



iotshow.in

January 11th-13th, 2016

Bangalore International Exhibition Centre (BIEC), Bengaluru

For details, Call: Jyotsna (011-40596605) or write to us at: efyexpo@efy.in

Co-located at

**INDIA
ELECTRONICS
WEEK**

January 11th-13th, 2016
Bengaluru

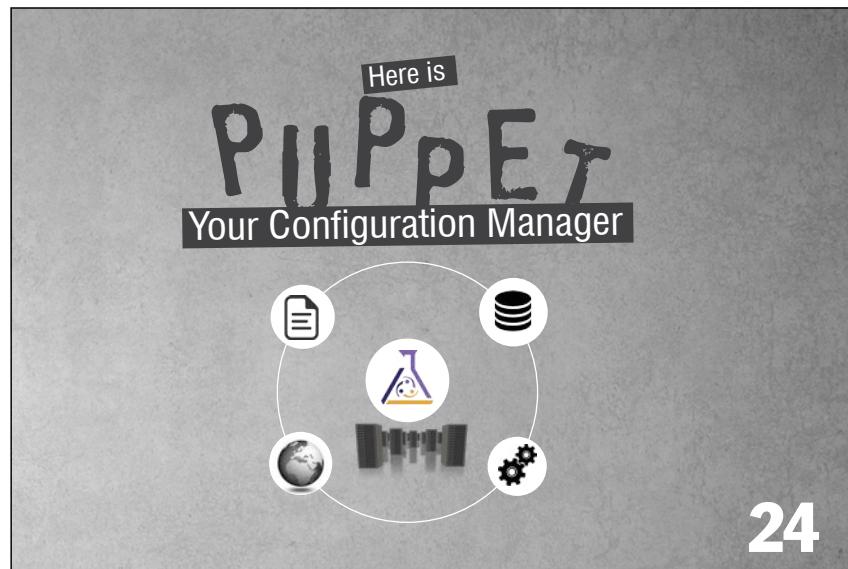
Contents

ADMIN

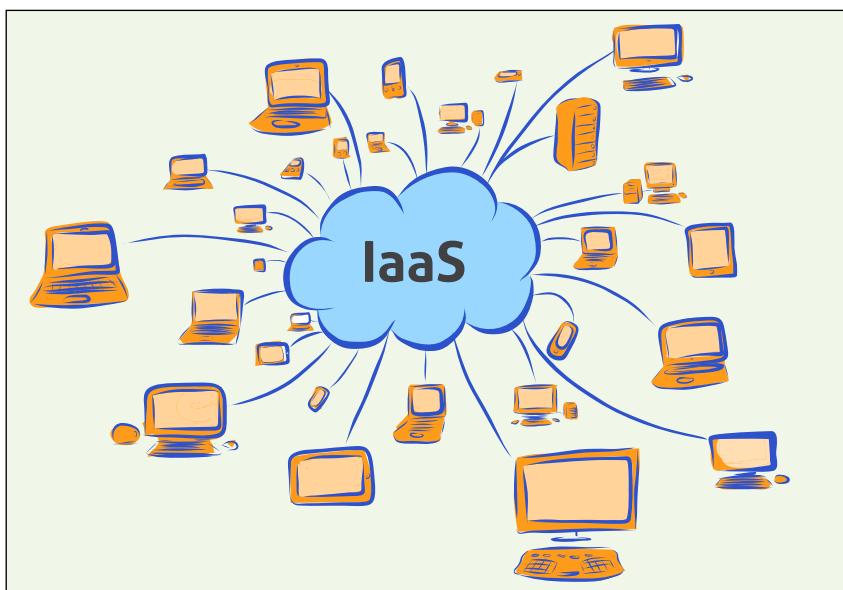
- 27** An Introduction to Distributed Replicated Block Device
- 31** Open Source Frameworks for Big Data Processing: An Overview
- 37** How MBaaS Providers are Simplifying App Development
- 43** NFV and the Service Provider Cloud
- 47** The Pros and Cons of Cloud Computing
- 54** Scalr: The Efficient Cloud Management Tool

DEVELOPERS

- 58** Achieve Shorter Software Delivery Timelines Using Cloud IDEs
- 62** An Introduction to Apache ActiveMQ
- 65** JavaScript The New Parts
- 68** The Importance of Memory Management
- 70** Building Manual and Automatic Counters in App Inventor 2
- 77** Haxe: Write Once, Convert to Many Languages
- 81** The Basics of MVC Architecture in PHP
- 83** A Sneak Peek into PHP 7



24



50 Introducing the Open Source IaaS Platform

REGULAR FEATURES

- | | |
|--------------------------|------------------------------|
| 10 New Products | 80 Editorial Calendar |
| 12 You Said It... | 104 Tips & Tricks |
| 13 FOSSBytes | |

Every Week
Free
Workshop

COURSES OFFERED

- ◆ Android
- ◆ C++
- ◆ Drupal
- ◆ Java
- ◆ Python
- ◆ Ruby on Rails
- ◆ PHP & MySQL
- ◆ OpenStack
- ◆ Linux Admin
- ◆ Hadoop Admin



ASSOCIATE PARTNER



COURSE SCHEDULE :

➤ Course Profile : *Linux & OpenStack System Administrator (1st Batch) & Android Developer(2nd Batch).*

➤ Online Written Test Fee : INR 100

➤ Last Date to Apply: 11th Dec 2015

➤ Online generation of Hall-Ticket: Till 12th Dec 2015

➤ Online Test: 13th Dec 2015 (Sunday)

➤ Results declared (Website) : 15th Dec 2015

➤ F2F Interview Test : 21st - 23rd Dec 2015

➤ Final Results declared (Website) : 26th Dec 2015

➤ Admission Start : 28th to 29th Dec 2015

Registration Start : 20th Oct 2015

Bank loan Assistance

Course Starting
4th Jan 2016

100% Job Guarantee With Through Bank

Foundation Course 5 days 40 Hrs	Advanced Course 5 days 40 Hrs	Embedded Course 5 days 40 Hrs	Foss Workshop Edge 30 Weeks 1000 Hrs	Foss Workshop Feature Course 30 Weeks 1000 Hrs (Only Fresher Can Apply)
--------------------------------------	------------------------------------	------------------------------------	---	--

Hands on Courses	Course Descriptions	Schedule of Courses (2015/16)
Linux & OpenStack Administrator	Duration: 30 Weeks (1000Hrs) Batch Size : 30 Eligibility: Freshers Engineering students (2015 & 2016), Selection: All India Entrance Test (13 th Dec 2015) 100% Job Guarantee (Otherwise 100% Fee Refund From Bank*)	Pune: 4 th Jan 2016
Android Developer	Duration: 30 Weeks (1000Hrs) Batch Size : 30 Eligibility: Freshers Engineering students (2015 & 2016), Selection: All India Entrance Test (13 th Dec 2015) 100% Job Guarantee (Otherwise 100% Fee Refund From Bank*)	Pune: 4 th Jan 2016
Drupal Expert	Duration: 30 Weeks (1000Hrs) Batch Size : 30 Eligibility: Engineering students, Working Engineering , Diploma holders , Hobbyists .	Pune: 4 th Jan 2016
Embedded Expert	Duration: 30 Weeks (1000Hrs) Batch Size : 30 Eligibility: Engineering students, Working Engineering , Diploma holders , Hobbyists .	Pune: 4 th Jan 2016
Foundation OpenStack	Duration: 5 Days (40Hrs) Batch Size : 25 Eligibility: Engineering students, working engineering, Diploma holders , Hobbyists .	Bengaluru 23 rd Nov to 27 th Nov 2015 Pune 7 th Dec to 11 th Dec 2015 Hyderabad 25 th Jan to 29 th Jan 2016

Hands on Workshop for Corporate

Hands on Workshop for Colleges & Institutes

Hire From Foss Workshop

Become Trainer

Limited Time Offer -
Buy 1 Study 2.

Hurry!

Few Seats Remaining.

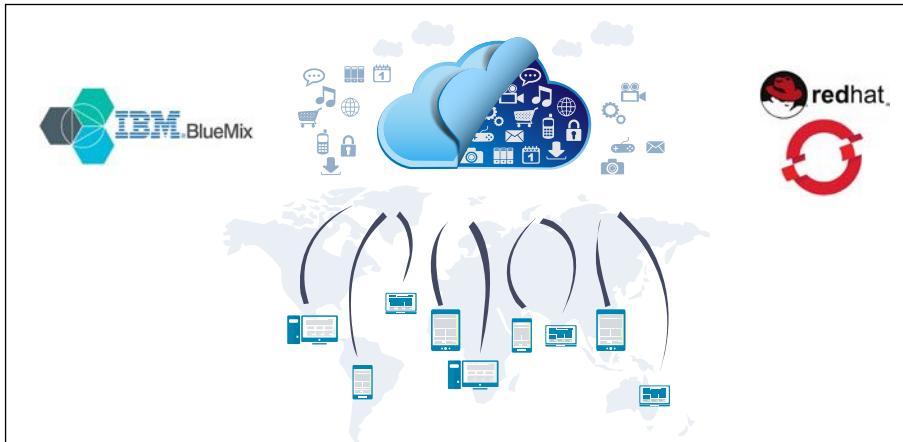
Salient Features

Inovative & Interactive Classes | Expert Trainers
24x7 Live Support | On Demand Shedule

Office Address: Foss Workshop Learning Pvt. Ltd. M-301, 2nd Floor Megacentre, Hadapsar, Pune-411028.

Contact us: (020) 6940 6940 **Visit us:** www.fossworkshop.com.

Contents



73 Creating Apps on Live Cloud Platforms: IBM Bluemix and Red Hat OpenShift



22 “Open source and a problem-solving attitude is the way to work at Zopper”

— Ashish Baweja, CTO, Zopper



101 “Open source technology plays a crucial role in the cloud business”
—Sunil Gupta, executive director and president, Netmagic

OPEN GURUS

94 A Brief on Search, Frames and Windows in GNU Emacs

96 Speech Signal Analysis Using Praat

FOR U & ME

85 Hybrid Drives: Get the Best of Both Worlds!

88 Contribute to SharedVehicles to Help Solve Urban Transportation Woes

91 What You Should Know about Open Source Licensing

COLUMNS

20 CodeSport

DVD Of The Month

Sample these Linux distros for geeks.

106

Editor
RAHUL CHOPRA

Editorial, Subscriptions & Advertising

DELHI (HQ)
D-87/1, Okhla Industrial Area, Phase I, New Delhi 110020
Ph: (011) 26810602, 26810603; Fax: 26817563
E-mail: info@efy.in

Missing Issues
E-mail: support@efy.in

BENGALURU
Ph: (080) 25260394, 25260023
E-mail: efyb@efy.in

Customer Care
E-mail: support@efy.in

Back Issues
Kits 'n' Spares
New Delhi 110020
Ph: (011) 26371661, 26371662
E-mail: info@kitnspareds.com

Advertising

CHENNAI
Ph: (044) 42994363
E-mail: efyenq@efy.in

HYDERABAD
Ph: (040) 67172633
E-mail: efyenq@efy.in

MUMBAI
Ph: (022) 24950047, 24928520
E-mail: efymum@efy.in

PUNE
Ph: 08800295610 / 09870682995
E-mail: efypune@efy.in

GUJARAT
Ph: (079) 61344948
E-mail: efyahd@efy.in

JAPAN
Tandem Inc., Ph: 81-3-3541-4166
E-mail: tandem@efy.in

SINGAPORE
Publicitas Singapore Pte Ltd
Ph: +65-6836 2272
E-mail: publicitas@efy.in

UNITED STATES
E& Tech Media
Ph: +1 860 536 6677
E-mail: veronique.lamarque@gmail.com

CHINA
Power Pioneer Group Inc.
Ph: (86 755) 83729797, (86) 13923802595
E-mail: powerpioneer@efy.in

TAIWAN
J.K. Media, Ph: 886-2-87726780 ext. 10
E-mail: jkmedia@efy.in

Exclusive News-stand Distributor (India)

IBH BOOKS AND MAGAZINES DISTRIBUTORS LTD
Unit No.10, Bezcela Complex, Opp. Suman Nagar,
Sion Trombay Road, Chembur,
Mumbai - 400 071
Phones : 022 – 0497401 / 02
E-mail: info@ibhworld.com

Printed, published and owned by Ramesh Chopra. Printed at Tara Art Printers Pvt Ltd, A-46,47, Sec-5, Noida, on 28th of the previous month, and published from D-87/1, Okhla Industrial Area, Phase I, New Delhi 110020. Copyright © 2015. All articles in this issue, except for interviews, verbatim quotes, or unless otherwise explicitly mentioned, will be released under Creative Commons Attribution-NonCommercial 3.0 Unported License a month after the date of publication. Refer to <http://creativecommons.org/licenses/by-nc/3.0/> for a copy of the licence. Although every effort is made to ensure accuracy, no responsibility whatsoever is taken for any loss due to publishing errors. Articles that cannot be used are returned to the authors if accompanied by a self-addressed and sufficiently stamped envelope. But no responsibility is taken for any loss or delay in returning the material. Disputes, if any, will be settled in a New Delhi court only.

SUBSCRIPTION RATES			
Year	Newstand Price ₹)	You Pay ₹)	Overseas
Five	6000	3600	—
Three	3600	2520	—
One	1200	960	US\$ 120

Kindly add ₹ 50/- for outside Delhi cheques.
Please send payments only in favour of **EFY Enterprises Pvt Ltd**. Non-receipt of copies may be reported to support@efy.in—do mention your subscription number.

Open Source Is Hot In The IT World

THE COMPLETE MAGAZINE
ON OPEN SOURCE

» To find dealers to buy copy from news-stand, visit: www.ezine.efymag.com

» To subscribe, visit: pay.efyindia.com

» To buy an ezine edition, visit: www.magzter.com choose Open Source For You



Hadoop
Apache
Postgres
Android
MySQL
OSS
PEARL
OpenStack
PHP
Drupal
Joomla

Get Set for the Biggest Celebration of Open Source Technologies – Open Source India 2015!

The stage is set for tech enthusiasts to come and celebrate the spirit of open source technologies. Open Source India (OSI) 2015 is just round the corner. Here's what this open source carnival will be all about...



The wait is over! It is that time of the year when open source enthusiasts and practitioners meet to celebrate the spirit of open source and take the technologies a notch higher. It is time for the 12th edition of Open Source India (OSI). This year's event is scheduled to take place at the NIMHANS Convention Center, Bengaluru, from November 19 to 20, 2015.

Started as LinuxAsia in 2004, OSI has been the most premier event bringing together the open source industry and the community. This year, the team behind the 12th edition of OSI aims to take this event a notch higher by focusing on the open source ecosystem in Asia, and more specifically, in India.

Last year, the two-day event witnessed a whopping 7500+ registrations, comprising IT decision-makers, developers and academia. OSI 2014 witnessed some interesting talks and discussions on the cloud, mobile app development, kernel development, databases, Web app development, et al. One of the most attractive features of the show was 'Success Stories', where CIOs and CTOs from companies like BookMyShow, Rediff.com, Goibibo.com, JustDial, Mahindra & Mahindra Finance, etc, came forward to share their open source success stories.

A glimpse of what's in store at Open Source India 2015!

The 12th edition of Open Source India will take the practice of 'industry meeting the community' a notch higher. Speaking

about the show, Rahul Chopra, editor, *Open Source For You* magazine, said, "At Open Source India, we try to bridge the gap between the industry and the community. Last year, the show witnessed some great interaction between industry stalwarts and community members. Such healthy discussions ended up in great technology innovations and the creation of job opportunities. We look forward to taking that journey a step further with Open Source India 2015."

OSI 2015 will feature a combination of exhilarating talks and hands-on workshops, during which techies will get an opportunity to learn and share their knowhow on various technologies.

With 10 tracks and eight workshops (as on date), the audience will get a good dose of the latest and the best in technology. The tracks at OSI this year include:

- Mobile App Dev Day
- Cloud Day
- IT Infrastructure Day
- Kernel Day
- Database Day
- FOSS for Everyone
- Web App Dev Day
- IT Implementation Success Stories
- OpenStack Mini Conference
- IoT and Open Source

Commenting on the tracks at OSI, Rahul Chopra said, "While we are touching upon some of the popular open

source domains like mobile app development, the kernel, databases, the cloud, Web app development and ‘FOSS for Everyone’, we are also introducing some of the real life open source success stories, which will be shared by the CIOs and CTOs of some of the biggest companies in the country. Another track to watch out for this year would be ‘IoT and Open Source’. In this track, the industry and community pioneers will share their experiences on how open source facilitates the revolution called the Internet of Things. Given the fact that the IoT revolution is gaining so much momentum, this track will surely attract a large number of industry and community techies.”

Yet another attraction is that OSI 2015 will have about 50-60 CIOs and CTOs from some of the biggest organisations of the country including Rediff, OLX, Flipkart, Mahindra Comviva, BookMyShow.com, AskMeBazaar, Max Life, HP, Zopper, Samsung, et al., participating in the event.

OSI has always believed in the philosophy of trying out new technologies hands-on, along with discussing them through talks delivered by experts. This year will be no different. The hands-on workshops have always been a super-hit amongst the techies. This year, Open Source India will host eight hands-on sessions on various technologies revolving around open source. The workshops at OSI 2015 are listed below.

Drupal 8 in a Day: You can get an excellent introduction to Drupal 8 in one day by attending this workshop. Learn about the most essential features and concepts of Drupal through hands-on activities. By the end of this day, you will be familiar with Drupal terminology and be able to identify how Drupal sites are constructed. You will also learn how to identify and choose modules to get the functionality you need.

Software Architecture - Principles, Patterns and Practices: If you are a developer or designer aspiring to become an architect and want to learn about the architecture of open source applications, or want to learn software architecture through case studies and examples, this workshop is certainly for you. It will introduce you to key topics in software architecture including architectural principles, constraints, non-functional requirements (NFRs), architectural styles and design patterns, viewpoints and perspectives, and architecture tools. A special feature of this workshop is that it covers examples and case studies from open source applications. What’s more, you’ll also get exposed to some free or open source tools used by practising software architects.

Big Data - What, Why and How: This workshop will help you to try your hand at installing the HADOOP ecosystem.

Shifting to Functional Programming - Lambda for Java Developers: Functional programming is on the rise. Almost all major mainstream languages support functional programming features, including C++, Java, Swift, Python and Visual Basic. With Java 8’s Lambda functions, Java now supports functional programming. Moving to functional

programming can result in significantly better code and productivity gains. However, it requires a paradigm shift – you need to move away from imperative and object-oriented thinking to start thinking functionally. That’s what this workshop will help you achieve. It will help you make the shift towards functional programming. The workshop will introduce Lambda functions in Java with examples from the Java library itself. Bring your laptops and you can get your hands dirty trying out the sample programs!



Cross-Platform Mobile Application Development - Consumer and Enterprise Application: Here’s a hybrid approach to developing cross-platform mobile business applications that uses open source technologies. The phenomenal growth of mobile devices has opened up avenues for organisations to integrate them into the mainstream computing environment. Today’s mobile applications deliver complex functionality on platforms that have limited resources for computing. Yet, unlike the PC-based environment, the mobile environment comprises a plethora of devices with diverse hardware and software configurations as well as communication intricacies. This diversity in mobile computing environments presents unique challenges in application development, quality assurance and deployment, requiring unique testing strategies.

The Cloud in a Day: Attend this workshop to build a cloud in just four hours with Apache CloudStack.

Design, Deploy and Manage APIs using the WSO2 API Management Platform: API is the new lingo of business development. A successful API strategy requires planning around usage scenarios, user profiles and ecosystems that your APIs are intended for. It also relies on a sound technology platform for reliable deployment and management of APIs. This workshop will bring together the strategic and technical domains of API design, and use the WSO2 API Management Platform to make concepts and best practices more understandable.

Tizen Web App Development – A Hands-on Session: The workshop will include an introduction to the Tizen SDK/IDE and tools, a hands-on session to develop a simple Web application, and a showcase of sample code usage in multimedia and location APIs. 

NEW PRODUCTS

SanDisk's wireless stick available in four variants

Global leader in flash memory storage solutions, SanDisk, has launched the 'connect wireless stick' in India. Available in four variants of 16GB, 32GB, 64GB and 128GB, the stick acts like a regular flash drive.

The wireless stick is compatible with devices like laptops, smartphones and PCs via its built-in Wi-Fi hotspot. It enables its users to wirelessly transfer and access bulky files, and stream videos to and from up to three mobile devices at the same time.



Price: ₹ 2,790, ₹ 3,790, ₹ 5,490 and ₹ 9,490 for 16GB, 32GB, 64GB and 128GB, respectively

The device is capable of streaming a single HD video at 2Mbps for up to 4.5 hours on a single charge. It works alongside the SanDisk Connect app, through which users can manage content between Windows, Android, iOS and Firefox devices. It is also equipped with features like auto-backup of images, videos, etc.

Address: SanDisk India, Survey No.143/1, Amani Bellandur Khane Village, Prestige Excelsior, Prestige Tech Park, Marathahalli-Sarjapur Outer Ring Road, Kadubeesanahalli, Varthur Hobli, Bengaluru – 560103;
Ph: +91-4242-2000

The third Tizen smartphone from Samsung is out

The South Korean electronics giant Samsung has recently unveiled its third Tizen based smartphone, the Samsung Z3, after discontinuing the Samsung Z and Z1.

The dual-SIM smartphone features a 12.7cm (5 inch) HD super AMOLED display with a resolution of 1280x720 pixels that delivers deeper colour reproduction.

The Z3 runs on Tizen 2.4 and is powered by a 1.3GHz quad-core Spreadtrum processor coupled with 1GB RAM.

It is equipped with 8GB storage expandable up to 128GB and a battery capacity of 2600mAh featuring 'ultra power saving' mode. The smartphone sports an 8 megapixel rear shooter with a F2.2 aperture and a 5 megapixel camera at the front, which is capable of shooting wider 120 degree selfies.

The device supports connectivity options of USB 2.0, Bluetooth 4.0, Wi-Fi and GPS.

The Samsung Z3 is available in gold, silver and black via Snapdeal and various retail stores.



Price: ₹ 8,490

Address: Samsung India Electronics Pvt Ltd, 2nd, 3rd and 4th Floors, Tower C, Vipul Tech Square, Golf Course Road, Gurgaon Sector 43, Gurgaon – 122002;

Ph: +(91)-124-4881234; **Website:** www.samsung.com

Micromax enters the power bank segment

Micromax has entered the portable power bank segment with its latest release called Yu Jyuice. The new power bank comes in 5,000mAh and 10,000mAh variants, which are useful for both light and heavy users. It sports an aluminium alloy shell and rounded edges with an 8mm, slim body.

The 5000mAh variant comes with one USB 2.0 port and one micro-USB port.

The device measures 139mm x 70mm x 7.9mm and weighs 250 grams. It delivers a conversion rate of up to 92 per cent. According to company claims, the 5000mAh charger can offer a 5V/2.1A output, and can charge a 2000mAh battery 2.3 times with a single charge.

The Jyuice 10,000mAh variant features two USB 2.0 ports, one micro-USB port and gives a maximum output of 5V/2.4A. The device measures 139mm x 70mm x 12.5mm and delivers a conversion rate of up to 92 per cent. According to the company, this variant of Jyuice can charge a 2000mAh device 4.6 times.

Both the variants feature an LED light, which turns red when the charge is below 30 per cent, yellow when between 30-70 per cent and green when above 70 per cent. The power banks are exclusively available via Snapdeal and other online portals.



Price: ₹ 699 and ₹ 1,099 for the 5000mAh and 10000mAh variants, respectively

Address: Micromax Informatics Limited, 90B, Sector-18, Gurgaon 122015;
Ph: +91-124-4811000; **Email:** info@micromaxinfo.com

A power bank from TP-Link that resembles a hip flask!



Price: ₹ 2,799

The global provider of networking products, TP-Link, has introduced a portable power bank called the TP-Link PB50. The device resembles the design of a hip flask, giving it a classy yet professional feel.

To ensure that your smartphone is charged all the time, the power bank is equipped with a 10,000mAh battery capacity.

The safety features of the power bank include:

1. Lithium polymer battery cells from LG, which offer durability and heat resistance.
2. The use of a porous PO material with high ventilation reduces the chances of short circuits inside the battery.
3. It is equipped with protection against over current, over discharge and over voltage to ensure that the smartphone is completely safe while charging.

Apart from the safety features, the good-looking PB50

power bank comes with dual USB 5V/1A and 5V/2A ports that allow charging of two different devices, simultaneously.

It is compatible with almost any smart device like the iPhone, iPad, Android smartphones, tablets or any other 5V input USB charged device. The device is designed with four green LED lights, in which each light indicates 25 per cent of the remaining battery charge.

The TP-Link PB50 is available online and at retail stores.

Address: TP-Link India Private Limited, Unit Nos 1-3, Ground Floor, Luthria House, Sativali Main Road, Vasai Road (East), Maharashtra – 401208; **Ph:** +91 9768012285; **Website:** www.tp-link.in

Lenovo's smartphone for camera lovers



Price: ₹ 25,499

The smartphone comes with inbuilt 32GB storage expandable up to 128GB via a microSD card. The VibeShot is powered by a 3000mAh battery measuring 142mm x 70mm x 7.3mm and weighs 145 grams. The connectivity options of Lenovo Vibeshot include 4G LTE, Wi-Fi, GPS/A-GPS, Bluetooth and micro-USB. The smartphone is available in three colours: crimson, pearl white and graphite grey via online and retail stores.

Address: Lenovo India Pvt Ltd, Vatika Business Park, 1st Floor, Badshahpur Road, Sector-49, Sohna Road, Gurgaon-122001;
Email: support@lenovoreg.com.

Chinese multinational computer technology company, Lenovo, has unveiled its latest smartphone that targets camera lovers. Called VibeShot, the smartphone is equipped with an 8 megapixel front camera and a 16 megapixel rear camera, with six-piece modular lens, BSI sensor and a tri-colour LED flash. It has a full frame 16:9 low light-sensor for shooting images in low light, dark interiors or at night.

As per company claims, the infrared autofocus feature in the smartphone focuses more accurately at speeds twice as fast as the normal auto focus.

Apart from the amazing camera, the smartphone features a 12.7cm (5 inch) full HD (1080x1920 pixels) display protected by Gorilla glass 3.0. It runs on Android 5.0 Lollipop and packs an octa-core 64-bit Snapdragon 615 SoC with 1.7GHz, along with 3GB of RAM.



Request for BOSS

It would be nice if you bundled the BOSS DVD with OSFY, or even BOSS-6, which is scheduled to be released shortly. This is 'India's own' (and possibly only) Linux OS version. I have been using it for some time and find it good. The developers, C-DAC, seem to have removed several irritants and bugs in Debian 7, its base system.

—J.S. Karkada,

eddykabootar@yahoo.com

ED: Thanks for writing to us. We will definitely try to include the BOSS (Bharat Operating System Solutions) OS in one of the upcoming DVDs. Please feel free to give us more suggestions and keep reading OSFY!



Have missed receiving my copy of OSFY

I have subscribed to OSFY since August 2015 for ₹1200 (with the 'Assured Gift With Every Subscription' offer). I received the August and September 2015 editions of the magazine in the first week of both months, respectively. But I have not received the October issue, as of the end of the second week of October. Please make arrangements to send the magazine as soon as possible.

—Kiran M.,

m.kiran02@gmail.com

ED: We sincerely regret the inconvenience caused. The magazine gets delivered latest by the 15th of every month. Though, generally it reaches our subscribers in the first week of every month. In case you do not receive your copy, you can send an email to support@efy.in with your concerns. They will be more than happy to resolve the issue and get the magazine delivered to you.



A helpful article

I began reading OSFY in my college library and love its contents as it has helped me expand my knowledge on open source. I was looking for a Web conferencing solution for an online course. The article in the June 2015 issue of OSFY on 'BigBlueButton: The Handy Web Conferencing Software' was really good and helped me a lot. Do include more such articles in future too.

—Rahul,

rahul88bhadwaj@gmail.com

ED: First of all, a big thank you for your words of appreciation. Through the years, we have constantly worked at improving the content, layout and design of the magazine in order to make your reading experience more enjoyable. We will certainly pass on your compliments to the author of the article. And as OSFY readers have always done, we urge you to continue sharing your opinion with us, as this motivates us and helps us to deliver even better.

I am a computer engineering student and saw the advertisement for an event on November 19 and 20 in Bengaluru. I believe this event is being conducted by OSFY. I wanted to know what it is all about and what I need to do to be able to attend it.

—Rajesh,

rajeshgope@gmail.com

ED: Yes, the event is being done by the EFY Group and OSFY is one of the publications of the Group. Open Source India (OSI) is a two-day event with multiple tracks and workshops running in parallel. It introduces you to a host of new open source technologies. It is being held at NIMHANS Convention Centre, Bengaluru. There are different levels of passes available, depending on what you want to attend. You can find more details in the curtain raiser published in this issue of OSFY or you can visit www.osidays.com. 



Please send your comments or suggestions to:

The Editor,
Open Source For You,
D-87/1, Okhla Industrial Area, Phase I,
New Delhi 110020, **Phone:** 011-26810601/02/03,
Fax: 011-26817563, **Email:** osfyedit@efy.in

Linux Mangaka moves from KDE to MATE

It seems that the Linux Mangaka developers have not been able to decide on the desktop environment for their distribution yet. The operating system, which initially used the GNOME based Pantheon desktop, had later moved to KDE4. Now, we hear that the distribution, backed by Animesoft International, wants to rely on MATE. Ubuntu based GNU/Linux distro, Mangaka Mou, has now entered the beta phase. Designed on low resources, Linux Mangaka Mou Beta is based on the 64-bit version of Ubuntu 14.04 LTS. This is a very important edition in the development cycle of the distro for Anime and Manga fans.

There has been no official announcement regarding the release of Linux Mangaka Mou, though some sites are offering links for downloads. However, as it is a pre-release version, it is not suitable for production use. Moreover, the version might include some unresolved issues which can affect the overall user experience.



Uptime Institute awards ESDS with Tier III certification

Uptime Institute has awarded ESDS, a managed cloud and data centre service provider, with a 'Tier III Certification of Design' for its upcoming Navi Mumbai data centre. ESDS is supposedly the first managed data centre service provider



in India to receive recognition from a global data centre authority like Uptime Institute. Considering the rapidly evolving IT scenario in the country, which is being further fuelled by the government's Digital India initiative, certified DC facilities such as ESDS will ensure reliability and availability for enterprise businesses. With their IT infrastructure housed in a Tier III DC, businesses can ensure availability of nothing less than 8754.4 hours

annually or 99.982 per cent availability.

It took ESDS 18 months to gain the certification, and its state-of-art data centre went live at Navi Mumbai in October, says one of the spokespersons from the firm. "ESDS has a very strong cloud platform, which has fault tolerance at the IT infrastructure level. Any kind of IT hardware failure cannot result in any downtime for the eNlight Cloud platform we have; but, having redundancy on the electrical infrastructure was extremely important for us. So we chose to get Uptime Institute's certification, as it has helped us to build a data centre where the electrical infrastructure will never fail, while the rest will be handled by our eNlight Cloud platform," says Piyush Somani, founder and MD, ESDS. Detailed information about the certification issued can be found at <http://uptimeinstitute.com/TierCertification>.

Ubuntu 15.10 to get thinner and handier scroll bars

Ubuntu 15.10 (Wily Werewolf) has been making headlines for its new features and fixes. If reports are to be believed, the upcoming operating system is all set to import scroll bars from upstream GNOME. The new scroll bars, which will be

SoftNAS introduces SoftNAS Cloud for Red Hat

SoftNAS Cloud, which continues to gain popularity on AWS Marketplace, now offers customers the reliability of Red Hat Enterprise Linux (RHEL), and the cost-effectiveness, scalability and flexibility of Amazon EC2. Red Hat and SoftNAS demonstrated SoftNAS Cloud for RHEL at AWS re:Invent 2015 in Las Vegas in the US, this October.

"Commercially supported solutions such as the SoftNAS Cloud are an important part of Red Hat's ecosystem aimed at accelerating customer adoption in the cloud. The expertise and storage software innovation that SoftNAS brings to our joint customers is a welcome addition to the 'Red Hat Connect for Technology Partners' programme," said Mike Werner, senior director, global technology partner programs, Red Hat.



Enterprises select RHEL for standardisation because of its enterprise class software hardening and quality of support. SoftNAS Cloud is now available on this leading Linux platform to help organisations gain additional storage agility and flexibility, while reducing storage costs compared to traditional storage hardware. Amazon EC2 running RHEL provides a dependable platform to deploy a broad range of applications.

Pricing for SoftNAS Cloud for RHEL starts at US\$ 0.18 per hour on AWS Marketplace.

Red Hat introduces open source curriculum

Red Hat Corporation has announced that its Red Hat Academy is available to universities with the required technology and support environments in the Europe, Middle East and Africa (EMEA) region. The academy offers an open source education programme that provides hands-on curriculum, labs, performance-based testing and educator training and support. Material for Red Hat Academy is available online. Participating universities are provided with access to Red Hat Academy curriculum enabling instructors to prepare for class instruction and personal certification. The academy also provides direct email support to instructors throughout the year for curriculum or systems assistance.



Instructors are encouraged to enrol for Red Hat's commercial training and exams for certifications such as Red Hat Certified System Administrator (RHCSA) or Red Hat Certified Engineer (RHCE).

The academy's Web based courses and exams are designed for instructor-facilitated classes in a lab setting, and instructors can choose course format, frequency and testing options that fit with their preferred teaching methods.

introduced in Ubuntu 15.10, are said to be thinner. This also reveals the fact that Ubuntu developers still rely a lot on GNOME.

Although the scroll bars are very thin, they are visible and more organised. "Well, by default (when no desktop scaling is in action) they are using 10px for input (while the drawn bar is 8px, but we use 2px extra for better precision). However, these sizes can be easily changed by just adding some .css code to your theme," explains Marco Trevisan, Ubuntu Unity developer. Ubuntu 15.10 (Wily Werewolf) is expected to see the light of day on October 22, 2016.

MapR introduces in-Hadoop document database

The first in-Hadoop document database will allow developers to quickly deliver scalable applications that also leverage continuous analytics on real-time data. With these major enhancements, developers benefit from the advantages of a document database combined with the scale, reliability and integrated analytics of enterprise-grade Hadoop and Spark. These capabilities will enhance an organisation's ability to improve a range of use cases—from personalising and delivering the best online shopping experience, reducing risk and preventing fraud in real-time, to improving manufacturing efficiencies and lowering costs.



To help developers experiment and experience MapR-DB, a developer preview with sample code is available for download.

The MapR distribution including Hadoop is architected to serve as a single platform for running analytics and operational applications. MapR-DB enables continuous analytics on real-time data, while reducing cluster sprawl, eliminating data silos, and lowering the TCO of data management.

The native JSON support in MapR-DB also enables faster time-to-value by letting developers quickly start up more business applications on more data types and sources.

MapR-DB supports the Open JSON Application Interface (OJAITM), which is designed to be a general-purpose JSON access layer across databases, file systems and message streams, enabling a flexible and unified interface to work with Big Data.

"MapR continues to build on the innovative data platform at the core of its Hadoop distribution," said Nik Rouda, senior analyst, Enterprise Strategy Group. "The addition of a document database capability (JSON) neatly extends the powerful NoSQL MapR-DB to seamlessly cover more types of unstructured business data. This makes it faster and easier to build Big Data applications, without the burden of shuffling data around first," he added.

"MapR enables enterprise developers to create new value-added applications, which they have only dreamed about so far," said Anil Gadre, senior vice president, product management, MapR Technologies.

End of the Ubuntu 15.10 development cycle

The Ubuntu 15.10 development cycle came to an end on October 8, when the operating system reached the final kernel freeze milestone. No more upgrades are being made past this point, and only bug fixes are permitted. The Ubuntu development cycle is not in sync with the Linux kernel, so the latest version released is not usually implemented. As it stands right now, the Linux kernel used in Ubuntu 15.10 is version 4.2. When a project hits one of these freeze milestones,

whether it's for software stacks or the kernel, no more upgrades are made. This allows developers to properly test everything and make sure that all is in order for the new launch.

Ubuntu usually ships with a specific branch of the kernel. Even if a new version of the kernel is released, Canonical doesn't usually upgrade it, at least not for the non-LTS versions of Ubuntu which are only supported for nine months. On the other hand, the Ubuntu LTS versions do get kernel upgrades with each new point release.

Google releases Chrome 46

Based on the open source Chromium 46 Web browser, Google Chrome 46 (the actual version number is 46.0.2490.71) comes as a free update for Windows, Mac and Linux users, and includes several security fixes as well as various under-the-hood improvements.

Unfortunately, Google didn't publish any details about the new features implemented in Google Chrome 46, so we can only assume that it's mainly a bug fix release that patches 24 security issues that have been discovered by various

developers.

"The Chrome team is delighted to announce the promotion of Chrome 46 to the stable channel for Windows, Mac and Linux," said Tina Zhou on behalf of the Google Chrome team and added, "Chrome 46.0.2490.71 contains a

number of fixes and improvements—a list of changes is available in the log."



NFV and SDN set for open source boost

NFV and SDN have formed a collaborative project under the umbrella of the Linux Foundation. The project will focus on using the work done by the ONOS (Open Network Operating System) community on open source platforms, solutions and ecosystem for service providers to monetise SDN/NFV, while helping vendors and service providers invent new business models. Specifically, the partnership will focus on creating SDN solutions tapping open source software platforms, white boxes, as well as network control and management applications to boost the creation and deployment of SDN platforms.

The Linux Foundation said it will assist ONOS to organise, grow and harness the open source community in taking its platform to the next level of production-readiness and drive adoption in production networks.

"Service providers are increasingly adopting open source software to build their networks and today are making open source and collaboration a strategic part of their business and an investment in the future," said Jim Zemlin, executive director of the Linux Foundation, in a statement.

"The Linux Foundation recognises the impact the ONOS project can have on service provider networks and will help advance ONOS to achieve its potential. The partnership combines the best of the two organisations' capabilities in support of a strategic vision to transform service provider infrastructure with open source SDN and NFV," he added.

Last month, the ONOS project released its fourth open source SDN platform under the name 'Drake', which is focused on infrastructure

HP announces launch of OpenSwitch community and NOS

HP and key supporters, Accton Technology Corporation, Arista, Broadcom, Intel and VMware, are delivering a community-based platform that provides developers and users the ability to accelerate innovation, avoid vendor lock-in and realise investment protection as they rapidly build data centre networks customised for unique business applications. The always-on nature of the global digital economy and the massive influx of data have changed how network operators, cloud providers, telcos and enterprises think about networking.



The volume of network traffic is growing exponentially, forcing large organisations to scale their data centres to new levels. These organisations need scale and flexibility to meet increasingly complex business and technology requirements.

"Open source software, with its collaborative community of specialised developers, accelerates innovation and improves the stability of the software platform, providing organisations with a more powerful infrastructure to support their specific business requirements," said Santanu Ghose, director, networking, HP India.

"The newly formed developer community and the Linux-based OpenSwitch NOS early release code will help developers address the rapidly evolving business and Web-scale networking needs in the industry," he added.

NetApp, Mirantis announce partnership

Mirantis has joined hands with NetApp and announced a partnership that combines the Mirantis' OpenStack with NetApp's storage infrastructure.

Kamesh Pemmaraju, vice president of product marketing, Mirantis, said that the joint reference architecture puts this expertise into the hands of the user. It leverages the latest release of Mirantis OpenStack to address high-availability use cases for both cloud-native and traditional applications targeted for high-SLA environments, he added.

One of the benefits that customers will get after this partnership is NetApp block storage drivers for MOS 7.0. This will enable enterprises to leverage NetApp's enterprise-class storage with the just-released Mirantis OpenStack 7.0.



A fuel plug-in for NetApp is also validated by Mirantis, which simplifies and automates deployment and configuration.

Soon, customers of the next version of OpenStack, called Mitaka, will be able to leverage NetApp Clustered Data ONTAP storage.

enhancements to support SDN and NFV 'case enablement'. The Drake release is said to include new security, configuration and application-level feature sets that support improvements to the northbound and southbound data flows.

New Linux kernel updates for Debian

There are four critical kernel vulnerabilities (CVE-2015-2925, CVE-2015-5257, CVE-2015-5283 and CVE-2015-7613) that have been addressed in the updates for Debian. The new kernel packages are 4.2.3-1 for Sid, 3.16.7-ckt11-1+deb8u5 for Jessie, and 3.2.68-1+deb7u5 for Wheezy. All the Debian GNU/Linux operating systems that are part of the branches mentioned above are vulnerable, which means that users are urged to update their Debian-based machines as soon as possible using the built-in package management utilities.

"Several vulnerabilities have been discovered in the Linux kernel that may lead to a privilege escalation, denial of service, unauthorised information disclosure or unauthorised information modification," said the Debian developers in the DSA-3372-1 security advisory.

The following security flaws have all been fixed in the new updates. The first security flaw was discovered by Jann Horn in the Linux kernel packages of Debian GNU/Linux 7 Wheezy. This could allow a privileged user with write permissions to access the entire file system when a sub-directory of the respective file system was bind-mounted into a mount namespace or chroot. This has been fixed for Debian GNU/Linux 8 Jessie systems.

The second kernel vulnerability was discovered by Moein Ghasemzadeh from the Istuary Innovation Labs. This is in the Linux kernel's USB drivers, as an attacker could use a USB device to crash the system by causing a denial-of-service attack by imitating a Whiteheat USB serial device.

The third flaw was discovered by Marcelo Ricardo Leitner in Debian GNU/Linux 8 Jessie's Linux kernel SCTP support, which could allow an attacker to crash the system by causing a denial-of-service attack by creating multiple SCTP sockets.

Lastly, the fourth kernel vulnerability was discovered by Dmitry Vyukov in Linux kernel's System V IPC implementation, which could allow a local user to gain root access, cause a denial-of-service attack, as well as gain access to or modify unauthorised information. This was possible because the System V IPC objects were made accessible before the initialisation of the ownership and several other attributes.

Nautilus' file search about to get big power-up

Nautilus' file search makes heavy use of GNOME's spiffy pop-over menus in an effort to offer a simpler way to narrow in on search results and find exactly what you're after. Developer Georges Stavracas is working on the new UI and describes the new editor as "cleaner, saner and more intuitive."

"Nautilus has very complex but powerful internals, which allows us to do many things. And indeed, there is code for the many options in there. So, why did it look so poorly implemented/broken?" he writes in his blog.

The new search filter interface surfaces many of the 'powerful internals' to the user. Searches can be filtered on the basis of content type,



name or by date range.

Changing anything in an app like Nautilus is likely to upset some users,



no matter how helpful and straightforward the new UI may seem. The outcry at the removal of ‘type ahead’ search in 2014 still rings loud in many ears.

GNOME 3.18, released last month, introduced a new file

progress dialogue box to Nautilus and better integration for remote shares, including Google Drive.

The reworked search UI is tentatively targeted for inclusion in GNOME 3.20, due in spring next year.

Microsoft launches Linux-powered OS

Principal architect at Azure Networking, Kamala Subramaniam, wrote on the company blog, “The Azure Cloud Switch (ACS) is our foray into building our own software for running network devices like switches. It is a cross-platform modular operating system for data centre networking built on Linux. ACS allows us to debug, fix and test software bugs much faster. It also allows us the flexibility to scale down the software and develop features that are required for our data centre and networking needs.”



In the blog post, Microsoft also acknowledged the value of open source, open standards and Linux, stating, “ACS believes in the power of open networking. ACS, together with the open, standardised SAI interface, allows us to exploit new hardware faster and enables us to ride the tide of ASIC innovation while simultaneously being able to operate on multiple platforms. Running on Linux, ACS is able to make use of its vibrant ecosystem. ACS allows people to use and extend open source, Microsoft, and third party applications.”

For now, it’s not known when Microsoft will release the source code of the project and which licence the company will use for it. The Linux kernel is licensed under GNU GPLv2, so it has to be a compatible licence.

Infosys, IBM join hands to develop cloud apps

IBM and Infosys have announced a collaboration in which Infosys will tap Bluemix, IBM’s cloud platform, to rapidly prototype, develop and deploy the next generation of cloud apps for its global client base.

The collaboration includes the launch of the Bluemix-powered Innovation Lab for the co-creation of apps with clients. It also includes training Infosys developers on Bluemix and cloud app development skills, as well as Infosys’ adoption of Bluemix dedicated to leverage its vast library of cognitive computing and analytics solutions.

Srikantan Moorthy, head of application development and maintenance, Infosys, said, “Digital transformation is unleashing the power of smarter processes for our clients. Our goal is to bring these advanced technologies to clients’ application landscape in the most rapid and collaborative way possible. Infosys will also incorporate Bluemix-related curriculum into its onboarding and training process.”

Steve Robinson, GM, IBM Cloud said, “Today a massive digital transformation is under way, underpinned by the disruptive forces of cognitive computing, analytics and IoT—all delivered through the cloud. Through Bluemix, developers can accelerate the deployment of these next-generation apps that leverage the most innovative advancements in technology. This collaboration with Infosys will further advance our joint clients’ journey and the adoption of Bluemix.”

Powered by Bluemix, the Infosys Innovation Lab will function as a place to innovate and co-create apps with a broad range of global clients. Staffed with a dedicated team of designers, Agile software specialists, and industry and technology architects, the lab is designed to rapidly deliver app prototypes.

IBM launched Bluemix with a US\$ 1 billion investment in 2014. It matured quickly to become one of the largest cloud foundry deployments in the world.



Are You In Suspense?

Are you wondering what happened to your order or payment, even after a fortnight? But did you mention your name and address clearly while sending payment, and what the payment was for?

We get several payments every month without adequate instructions or even identity of the person sending the payment. As a result we are unable to take appropriate action on these, till we get a complaint - if at all.

At EFY, we want you to be our satisfied customer. So please do let us know in case your payment or instructions are not even acknowledged within a fortnight. We are not that lazy!

Email: support@efy.in
Ph: 011-26810601 / 02 / 03 extn: 202

EFY GROUP
Technology Drives Us

UK government switches from MS Office to LibreOffice

In an effort to shed its dependency on proprietary software, the UK government has entered into a new commercial deal with a British open source software company, Collabora Productivity that adapts LibreOffice for use in enterprises.

Such a high profile move by the UK government, through its Crown Commercial Service (CCS) body, is bound to turn the spotlight on Collabora, one of the biggest contributors to LibreOffice. It has the largest team of certified LibreOffice engineers.



involves more than just a simple adoption.

The deal also includes a product named CloudSuite, which provides its users with cloud based and mobile access, and the ability to view and edit locally on devices as well as online in Web browsers. CloudSuite will be released soon and is going to bring crucial functionality.

Interestingly, it is getting difficult for Microsoft to keep pace with a free office suite. Collabora is not selling the UK government anything, but has been hired to provide technical assistance and support for a fee that will be a lot cheaper than paying for support and licences, at the same time.

The deal between the Crown Commercial Service (CCS), an executive arm of the Cabinet Office, and Collabora Productivity underscores the government's commitment to open source and the Open Document Format, and offers major cost savings for public sector bodies, an official announcement stated.

It is likely that we may get to see many more such switches happening next year, with other governments, giving LibreOffice more reach.

ownCloud Server 8.2 officially released

One of the world's most popular free and open source self-hosting cloud solutions, ownCloud Server 8.2 is the second point release in the 8.x series. It is a feature release that introduces numerous design updates like a redesigned gallery app, a brand new sidebar, support for notifications and better handling of mime types. It also has tons of security improvements. Additionally, ownCloud Server 8.2 updates introduce a policy for retaining file versions and files that have been deleted in the build-in text editor, add encrypt and decrypt commands, and greatly update the ownCloud API (application programming interface) with Query Builder, Capabilities Manager and Modular Authentication support.

The ownCloud Server 8.1 previous release was all about stability, performance and security. Now, more work has been put into the user experience, said Jos Poortvliet, community manager at ownCloud.

In the history of the project, ownCloud Server 8.2 might not be the biggest update, but it is a very important one. It adds many important changes that affect the overall user experience. The new ownCloud release introduces a new appliance that contains the ownCloud Proxy app announced in September this year, in addition to the new features mentioned above.

For users, upgrade instructions are provided on the project's website, along with any other information needed to start deploying your own cloud!



PIONEERS IN TECHNOLOGY. PARTNERS IN SUCCESS.



Mr Kapur, Mr Ajwani and Mr Verma are not the only ones who owe their success to *Electronics For You*. The technology magazine that started off in 1969 is today South Asia's most popular one. Covering the latest on emerging technologies, changing industry trends, evolving trade practices...

Not surprisingly, top decision makers and those high up on the corporate ladder in the electronics fraternity treat EFY as their Bible.

- Over half a million readers • India's largest selling technology magazine • Caters to the entire electronics fraternity

CODE SPORT



In this month's column, we feature a set of computer science interview questions.

Last month, we had discussed a set of questions on data management and cloud computing. We continue our discussion of computer science questions this month, focusing on machine learning and natural language processing. Let us get started with some tough questions to whet your appetite.

1. In natural language processing, the last couple of years have seen word embedding becoming extremely popular in various NLP tasks such as named entity recognition, sentiment analysis, semantic similarity detection, etc. Most of you would be familiar with Google's word2vec package (<https://code.google.com/p/word2vec/>). What are word embeddings? Can you explain briefly the two major algorithms in word2vec - namely the skip-gram method and continuous bag of words (CBOW) method? Here is a link to an excellent article which provides much of the historical background on word embeddings: <http://gavagai.se/blog/2015/09/30/a-brief-history-of-word-embeddings/>.
2. Over the last few years, we have also seen the rise of deep learning methods in computer vision, pattern recognition, etc. While deep learning achieved considerable popularity with Google's work (<http://www.wired.com/2012/06/google-x-neural-network/>) on using deep learning algorithms to automatically label CAT images, there are still open questions on why deep learning is effective in computer vision, speech recognition and various other pattern-recognition tasks. Can you explain how deep learning is different from traditional methods of machine learning? If you are interested in knowing more about deep learning and how to code on it, I would like to suggest the online tutorial by Andrej Karpathy titled, 'Hacker's Guide to Neural Networks'.
3. Can you explain what a kernel function is, in machine learning? How is it different from a

standard ML classifier, which is based on features?

4. What are 'Support Vector Machines (SVMs)'? Let us assume that you are given a set of data points, some marked red and some blue. You are asked to write a classifier, which can separate the blue and red points. You have written your favourite SVM based classifier for this task. Now you are given Figure 1 as input.

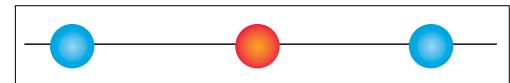


Figure 1: [SVM]

What is the separation boundary your classifier would find for this input? Can you use an SVM based classifier for separation for this input? If yes, explain how. If not, explain why.

5. What is a convex function? What is the significance of convex functions in machine learning algorithms?
6. In supervised machine learning problems, what is the process of cross-validation? How does cross-validation help prevent a model from being overfitted to training data?
7. In machine learning, can you explain the difference between model parameters and hyper-parameters? Do all models have hyper-parameters? Can you explain what would be the hyper-parameter for the standard linear regression model of the following form:

$$w^T x = y$$

...where x is the feature vector for the data, y is the scalar variable that needs to be predicted and w is the model parameter's matrix?

8. You are running a randomised clinical trial, where you are asked to determine whether a newly discovered drug for cancer is effective. You have a control group of patients who do not get the new

- drug and a target group of patients who get the drug. Let us assume that the clinical effectiveness of the new drug is measured in terms of the reduction in the number of deaths in the control and target groups for six months following the start of the clinical trial. Can you apply the concept of A/B testing for this situation? If not, explain why.
9. Consider the same example as above. You find that 20 people out of 400 died in the control group during the trial period, whereas 17 people out of 400 people died in the target group (who received the new drug). Does this establish the clinical effectiveness of the new drug against cancer? Can you explain your answer in terms of the statistical significance of the results?
 10. Let's consider Question (8) again. While Questions (8) and (9) make the implicit assumption that the difference in the number of deaths seen in the control and target groups is due to the new drug being administered to the target group, is this a correct assumption to make? If yes, substantiate why. If not, give an example to explain your answer.
 11. What are the confounding factors in statistical experiments? How does one account for confounding factors when conducting randomised experiments?
 12. A well-known phrase used in machine learning and statistics is: "Correlation is not the same as causation." Can you give an example to illustrate this statement?
 13. There are three kinds of analytics -- prescriptive analytics, descriptive analytics and predictive analytics. Can you explain each of these terms and describe how they are different from each other?
 14. There are a number of data analytics frameworks which can scale to analyse vast amounts of Web-scale data. Most of you would be familiar with Hadoop frameworks. Of late, there has been considerable interest in the Apache Spark framework as an alternate to Hadoop. Can you explain the major differences between the Hadoop and Apache Spark platforms? What would be your preferred choice for running a purely batch analytics application? And what would be your choice for running a real-time analytics application?
 15. Most of you would be familiar with recommender systems, which can be built using either collaborative filtering approaches or knowledge based recommendations. Can you explain the difference between the two approaches?
 16. What is the 'cold start' problem in recommender systems? How is it handled in collaborative filtering systems? Does this problem occur in knowledge based recommender systems?
 17. We are all familiar with search engines displaying ads that are related to our search query on the Web pages we are browsing. For example, if you were searching for information on Samsung mobile phones, your search engine will display ads for different mobile phones and mobile phone accessories. Can you explain how the search engines decide which ads to display for each user? For example, consider two different users, A and B, who are both looking for the same product. Are they likely to be displayed the same set of ads by the search engine? If you think not, explain how the search engines would decide what ads to display for each user, even though their search query term was identical.
 18. Given that one word can have different meanings (this is known as polysemy), how do information retrieval systems figure out what is the specific information that is being searched for? For example, consider a user who searches for 'windows'. How would the search engine determine whether he is looking for Microsoft Windows or glass windows? Blindly displaying search results for both meanings is likely to impact the user experience. If you are the designer of the search engine, explain how you would attack this problem? Consider the following related problem. When the user types in the search query as 'automobile', how does the search engine include Web pages that talk about cars as well?
 19. Consider the display of search query results by search engines. Given a search query term, how do search engines rank the relevant results? What factors are taken into account in ranking the search results? What is meant by 'filter bubble' in search engines and how would you avoid the problem?
 20. While traditionally, search engines have been tasked with the goal of returning Web pages relevant to the search query, they are increasingly being used to answer user questions. Google uses its knowledge graph to provide fact checked information about well-known personalities along with traditional search results. For example, if you type the question, "What is the first book by J.K. Rowling?" you will get a text box containing the relevant snippet as the answer along with standard search results. How would you design a search engine that can provide direct answers to questions?
- If you have any favourite programming questions/software topics that you would like to discuss on this forum, please send them to me, along with your solutions and feedback, at sandyasm_AT_yahoo_DOT_com. Till we meet again next month, happy programming!
- END** 

By: Sandya Mannarswamy

The author is an expert in systems software and is currently working as a research scientist at Xerox India Research Centre. Her interests include compilers, programming languages, file systems and natural language processing. If you are preparing for systems software interviews, you may find it useful to visit Sandya's LinkedIn group *Computer Science Interview Training India* at <http://www.linkedin.com/groups?home=&gid=2339182> & [HYPERLINK "http://www.linkedin.com/groups?home=&gid=2339182"](http://www.linkedin.com/groups?home=&gid=2339182) gid=2339182



Zopper: Looking For Problem Solvers and Open Source Believers

The Zopper app on your mobile is like a marketplace that you could visit from the comfort of your home or office, sans the traffic snarls and the parking woes. **Ashish Baweja, CTO, Zopper**, who has been experimenting and working with open source technologies for the past 17-18 years, has used his experience in this domain to help build Zopper completely on open source software. Here, he tells **Diksha P. Gupta** from **Open Source For You**, why he has done so, and why he is so fond of open source software.

Zopper, the hyper-local mobile marketplace, recently unveiled its revamped app. The latest version of the app features several improvements on the existing framework for additional user convenience. The Zopper app now allows users to search for products by model number, category or brand to find out the best local prices, while on the go, from nearby trusted stores -- all in the blink of an eye. The app even allows users to request for price quotes from nearby merchants who usually respond within minutes of getting a lead. It can also be personalised by entering a few details, which would then allow it to recommend products based on customer preferences and shopping

patterns. Users also have the option of buying from the app itself, avail cash on delivery, or even visit the nearby store to experience the product prior to buying it.

What is exciting for the readers of *Open Source For You* is that the Zopper app is completely based on open source technologies. Ashish Baweja, CTO, Zopper, is proud of the fact that he is an open source freak. After all, he gets to explore the “best of the tech world.”

Open source is the way Zopper works

When asked whether he uses open source technologies at Zopper, Baweja candidly responds, “A lot! Our core platform is in Python. We use Django and Falcom

frameworks for the database. We are using MySQL, Redis, MongoDB, et al. Everything we do is essentially open source based.”

Zopper’s CTO believes that open source technologies can do it all. He says, “I have been using open source for over 17-18 years, right since my college days when we used to get Linux on floppies. My encounter with open source technology began with some Linux CDs provided with magazines. Even during my college days, I started a few small ventures based on open source technologies. At that time, Perl was a popular language. I used to experiment with code and that was a whole lot of fun. This ‘fun’ aspect of Linux has been the driving force behind my passion for open source technologies. And I continue having fun in my career too. Prior to my job at Zopper, I was a technical consultant with many companies, where I propagated open source technologies. I also worked with Macmillan, where I switched a lot of their processes to open source.”

It's just about the innumerable advantages...

When you talk to an open source fan like Baweja, you know you will get nothing but the positives about this freedom-friendly technology. He has his reasons for banking on open source technology. He shares, “The best thing is that open source technology allows access to the code, allowing techies like us to have full control of the system. Open source technologies, of course, are free and affordable. They are essentially driven by an entire community and that is a big advantage. Using open source technologies allows us to be a part of that entire community’s ecosystem, where other tech majors are also contributing. With so many players being a part of a product’s development, the robustness of the product increases manifold.”

The other advantage is the scalability he gets with open source technology. He says, “Zopper is India’s largest hyper-local mobile marketplace featuring over 500,000 offline retailers. Founded by serial entrepreneurs Surjendu Kuila and Neeraj Jain, Zopper taps the benefits of shopping at a retail store, and combines them with the convenience of shopping via mobile phones. Users can buy from the comfort of their homes/offices and get the product instantly from their nearby trusted stores. And in all these efforts, technology is at the forefront. If we need to add more and more retailers and customers, we need technology that allows scalability. With

open source, you definitely have the bandwidth to scale as much as you want.”

The community is really important...

About interacting with the community, Baweja says, “Yes, we do that a lot. We try and participate in every possible way. While we participate in events like PyCon and other similar conferences, we also closely interact with the communities around Python, Django and other such technologies that we use. As far as our contributions go, though we have contributed some bits and pieces to the community in the past, it’s not as much as we would like to do. That is definitely going to be one of our focus areas in the future.”

Want a job at Zopper? Update your GitHub profile

Zopper houses an engineering team of about 40 people. And the company has a very unique way of hiring too. Baweja says, “Our standard hiring process involves solving a problem and submitting it in GitHub. We look at candidates’ GitHub profiles and evaluate the kind of work they have done. Essentially, we look for problem solvers. Knowledge of a specific technology is not all that important. What we look at is the right attitude to handle a problem and a situation. In a nutshell, we look at the core engineering skills and the problem-solving attitude. I believe that if someone is good with Java, they can also work with Python or any other programming language. Learning a programming language is simple, if someone has the problem-solving attitude. Open source technologies contribute a lot in that direction. I know people who randomly keep installing open source technologies that are available, and try them out and figure out what all can be done with them. So it is the ‘other way round’ approach. Instead of looking for a solution, we look for what best can be done with a piece of software and then fit it as a solution. I believe that people who do that, have a much better approach towards problem solving.”

The disadvantages are non-existent

Baweja concludes, “from Zopper’s perspective, I don’t think we have faced any disadvantage by using open source technologies. However, larger corporates may face some bottlenecks if they use open source. They may want instant solutions, and not have all the time to explore the options and solve the problems in the open source domain. We handle such issues ourselves, so this is not a disadvantage for us.”



Here is PUPPET

Your Configuration Manager



Puppet is an open source configuration management utility that runs on several UNIX-like systems as well as Microsoft Windows. It helps sysadmins automate many repetitive tasks.

A group of computers in a network is called a cluster. In many organisations, you will find the server being set up in a cluster so that if one machine fails, the process goes on. Even in Hadoop, the data is scattered over the cluster. Maintenance of these clusters is a cumbersome job for the administrator and developers. For example, let's assume that in a cluster of 1000 nodes (nodes refer to physical machines), there is a software update that needs to be applied on all the machines. Now, if you don't have a configuration management system, then you have to manually install the software on all the machines. This is really tiring when you have many software updates or configuration changes on a frequent basis.

To ease the job, configuration management systems were created and Puppet is one of them. It allows us to define the state of our IT infrastructure, and then automatically enforces the desired state. Puppet automates every step of the software delivery process, from provisioning of physical and virtual machines to orchestration and reporting; from early-stage code development through testing, production release and updates. Puppet has two solutions - one for configuration management and the other for data centre automation. In this article, we will look at its configuration management feature.

Puppet works in a master-slave design, where all the required configuration is done on the master and is reflected in all the slaves on a regular basis. The slave machine syncs up with the master every 30 minutes to reflect the configuration of the master. I will tell you how to: a) Install the latest version of Ruby; b) Install and configure a central Puppet master server; c) Install and configure a Puppet agent node; d) Sign the agent's SSL certificate request on the master, and finally; e) Execute a simple test with Puppet.

Ruby is one of the basic requirements for installing and configuring the Puppet agent/master set-up. So let's install it in on both master and slave. Use one of the following versions of MRI (standard) Ruby: 2.1.x, 2.0.x, 1.9.3 or 1.8.7.

Steps for installing Ruby

1. We first need to install all the required packages for Ruby installation on our system using the following commands:

```
yum install gcc-c++ patch readline readline-devel zlib zlib-devel
yum install libyaml-devel libffi-devel openssl-devel make
yum install bzip2 autoconf automake libtool bison iconv-devel
```

2. Download the latest Ruby source file using the following link: <https://www.ruby-lang.org/en/downloads/>
3. Extract the source file using the commands given below:

```
gunzip <sourcefile.tar.gz>
tar -xvf <sourcefile.tar>
```

To configure and install Ruby 2, type:

```
cd <sourcefile>
./configure
Make
make install
```

4. Now that the latest version of Ruby has been installed successfully in our server, we can verify the installed Ruby version by executing the command given below:

```
ruby -v
```

After Ruby is installed on the master and slave, let's configure Puppet on the master. The newest version of Puppet can be installed from the yum.puppetlabs.com package repository. To enable the repository, run any one of the following commands that correspond to your OS version.

For, Enterprise Linux 7, type:

```
$ rpm -ivh https://yum.puppetlabs.com/puppetlabs-release-el-7.noarch.rpm
```

For Enterprise Linux 6, type:

```
$ rpm -ivh https://yum.puppetlabs.com/puppetlabs-release-el-6.noarch.rpm
```

For Enterprise Linux 5, type:

```
$ rpm -ivh https://yum.puppetlabs.com/puppetlabs-release-el-5.noarch.rpm
```

Now run the command below which will install the Puppet master:

```
yum install puppet-server
```

Don't start the Puppet master service now. For now, delete any existing SSL certificates that were created during the package installation using the following command:

```
rm -rf /var/lib/puppet/ssl
```

Edit the master's *puppet.conf* file with the following content:

```
Method:    O (0 undocumented)
Total:    O (0 undocumented)
0.00% documented

Elapsed: 0.0s

config.status: creating x86_64-linux-fake.rb
./miniruby -I. -I.. -I.ext/common ..../tool/runruby.rb --extout=.ext -- --d
-ext" --mflags="" --make-flags="" --data-mode=0644 --prog-mode=0755 --instal
installing binary commands: /usr/local/bin
installing base libraries: /usr/local/lib
installing arch files: /usr/local/lib/ruby/2.2.0/x86_64-linux
installing pkgconfig data: /usr/local/lib/pkgconfig
installing command scripts: /usr/local/bin
installing library scripts: /usr/local/lib/ruby/2.2.0
installing common headers: /usr/local/include/ruby-2.2.0
installing manpages: /usr/local/share/man/man1
installing extension objects: /usr/local/lib/ruby/2.2.0/x86_64-linux
installing extension objects: /usr/local/lib/ruby/site_ruby/2.2.0/x86_64-linu
installing extension objects: /usr/local/lib/ruby/vendor_ruby/2.2.0/x86_64-li
installing extension headers: /usr/local/include/ruby-2.2.0/x86_64-linux
installing extension scripts: /usr/local/lib/ruby/2.2.0
installing extension scripts: /usr/local/lib/ruby/site_ruby/2.2.0
installing extension scripts: /usr/local/lib/ruby/vendor_ruby/2.2.0
installing extension headers: /usr/local/include/ruby-2.2.0/ruby
installing default gems: /usr/local/lib/ruby/gems/2.2.0 (build_info, cdc
bigdecimal 1.2.6
io-console 0.4.3
json 1.8.1
psych 2.0.8
rake 10.4.2
rdoc 4.2.0
installing bundle gems: /usr/local/lib/ruby/gems/2.2.0 (build_info, cdc
power_assert-0.2.2.gem
minitest-5.4.3.gem
test-unit-3.0.8.gem
installing rdoc: /usr/local/share/ruby/2.2.0/system
installing rdoc-docs: /usr/local/share/doc/ruby
[root@bac-lnx-vm50 ruby-2.2.2]#
[root@bac-lnx-vm50 ruby-2.2.2]#
[root@bac-lnx-vm50 ruby-2.2.2]# ruby -v
ruby 2.2.2p95 (2015-04-13 revision 50295) [x86_64-linux]
[root@bac-lnx-vm50 ruby-2.2.2]#
```

Figure 1: Installed Ruby version

```
[main]
logdir=/var/log/puppet
vardir=/var/lib/puppet
ssldir=/var/lib/puppet/ssl
rundir=/var/run/puppet
factpath=$vardir/lib/facter
```

```
[master]
```

These are needed when the Puppetmaster is run by passenger and can safely be removed if WEBrick is used.

```
ssl_client_header = SSL_CLIENT_S_DN
ssl_client_verify_header = SSL_CLIENT_VERIFY
```

Now include the DNS name at which agent nodes can contact the master. Update the master's *puppet.conf* [main] section with the following two lines:

```
Surname = puppet
dns_alt_names = puppet,<Puppet Master's Hostname or FQDN>
```

Create the CA certificate by running the following command:

```
puppet master --verbose --no-daemonize
```

Verify the certificate information that was just created with the following command:

```
puppet cert list -all
```

```

login as: root
root@bac-lnx-vm105's password:
Last login: Wed Aug 19 01:33:56 2015 from 10.65.200.241
MOTD for Thu May 21 10:22:06 EDT 2015
-----
# HOST NAME: bac-lnx-vm105          ID: S10aed59
# Kickstart Date:      Thu 21 May 2015 09:57:33 AM EDT
# OS Version:          RHEL6.6 64-bit
#
# vCPU:                  2 cores
# CPU Chips:            Intel ixDualCore 3.07GHz 64-bit
#
# RAM:                  4 GB
# Swap:                 4 GB
#
# Disk Drive Info
/dev/sda:              59.1 GB
#
# Note: This MOTD file gets updated automatically.
#       Changes to this file will be automatically overwritten!
#       Please open a WRT case if you have any problems with this machine
#
[root@bac-lnx-vm105 ~]# cd /etc/puppet/puppet.conf
[bash: cd: /etc/puppet/puppet.conf: Not a directory]
[root@bac-lnx-vm105 ~]# /etc/init.d/puppet start
Starting puppet agent:                                         [ OK ]
[root@bac-lnx-vm105 ~]# puppet agent --test
Info: Retrieving pluginfacts
Info: Retrieving plugin
Info: Caching catalog for bac-lnx-vm105.cisco.com
Info: Applying configuration version '1439965225'
Notice: /Stage[main]/Main/File[/tmp/example-ip]/ensure: created
Notice: Finished catalog run in 0.04 seconds
[root@bac-lnx-vm105 ~]# cat /tmp/example-ip
Here is my Public IP Address: 10.81.89.237.
[root@bac-lnx-vm105 ~]# 
[root@bac-lnx-vm105 ~]# 
[root@bac-lnx-vm105 ~]# 

```

Figure 2: Result

Now create the manifest file of Puppet master touch, /etc/puppet/manifests/site.pp. After all this, start the Puppet master service with the following command:

```
/etc/init.d/puppetmaster start
```

Once Puppet is installed on the master, let's move on and customise it on the Puppet slave. Ruby must be installed on the slave now; so copy the repository package from the Puppet master server and enable the repository with the following command:

```
rpm -ivh puppetlabs-release-el-6.noarch.rpm
```

Now type `yum install puppet` to install the Puppet agent. After doing this, update the `puppet.conf[agent]` with the following line in quotes: “`server = <Puppet Master's Hostname or FQDN>`”. Now start the Puppet agent service using the following command:

```
$ /etc/init.d/puppet start.
```

After successfully installing the Puppet on the master and client, we have to sign the SSL certificate on the master. The first time Puppet runs on an agent node, it will send a certificate-signing request to the Puppet master. Before the master is able to communicate and control the agent node, it must sign that particular agent node's certificate. On the Puppet master, run the following command to list all unsigned certificate requests:

```
$ puppet cert list
+ "<Agent's Hostname or FQDN>" (SHA256) 21:EF:ED:81:95:F0:ED:6D
: A5:1D:05:48:C9:01:5D:A6:FE:0C:FB:A4:89:FF:45:CF:26:53:1C:BB:1F
: 6D:1E:CC
```

On the Puppet master, run the command given below to sign the agent's certificate request:

```
$ puppet cert sign <Agent's Hostname or FQDN>
```

The Puppet master can now communicate and control the agent to which the signed certificate belongs. Now that everything is up and running, let's execute a simple test with Puppet. In the manifest file of the master, write some code to show the IP address.

Edit the main manifest file (`/etc/puppet/manifests/site.pp`) on the Puppet master with the following content, and then save it:

```
file {'/tmp/example-ip', # resource type file and filename
ensure => present, # make sure it exists
mode => 0644, # file permissions
content => "Here is my Public IP Address: ${ipaddress_eth0}.\n",
# note the ipaddress_eth0 fact
}
```

The Puppet agent periodically checks in with the Puppet master (typically every 30 minutes). It is also possible to initiate the check for a particular agent node manually, by running the following command (on the agent node):

```
$ puppet agent -test
```

Then run the following command to print the file:

```
$ cat /tmp/example-ip
```

The output looks like what's shown below (with that node's IP address):

```
Here is my Public IP Address: 10.81.89.101.
```

Thus we can configure the master and this configuration will be reflected in slave nodes. The benefits of this tool are:

- a) It frees up time to work on projects that deliver more business value.
- b) It ensures consistency, reliability and stability.
- c) It facilitates closer collaboration between sysadmins and developers, enabling more efficient delivery of clean code that delivers real business value.

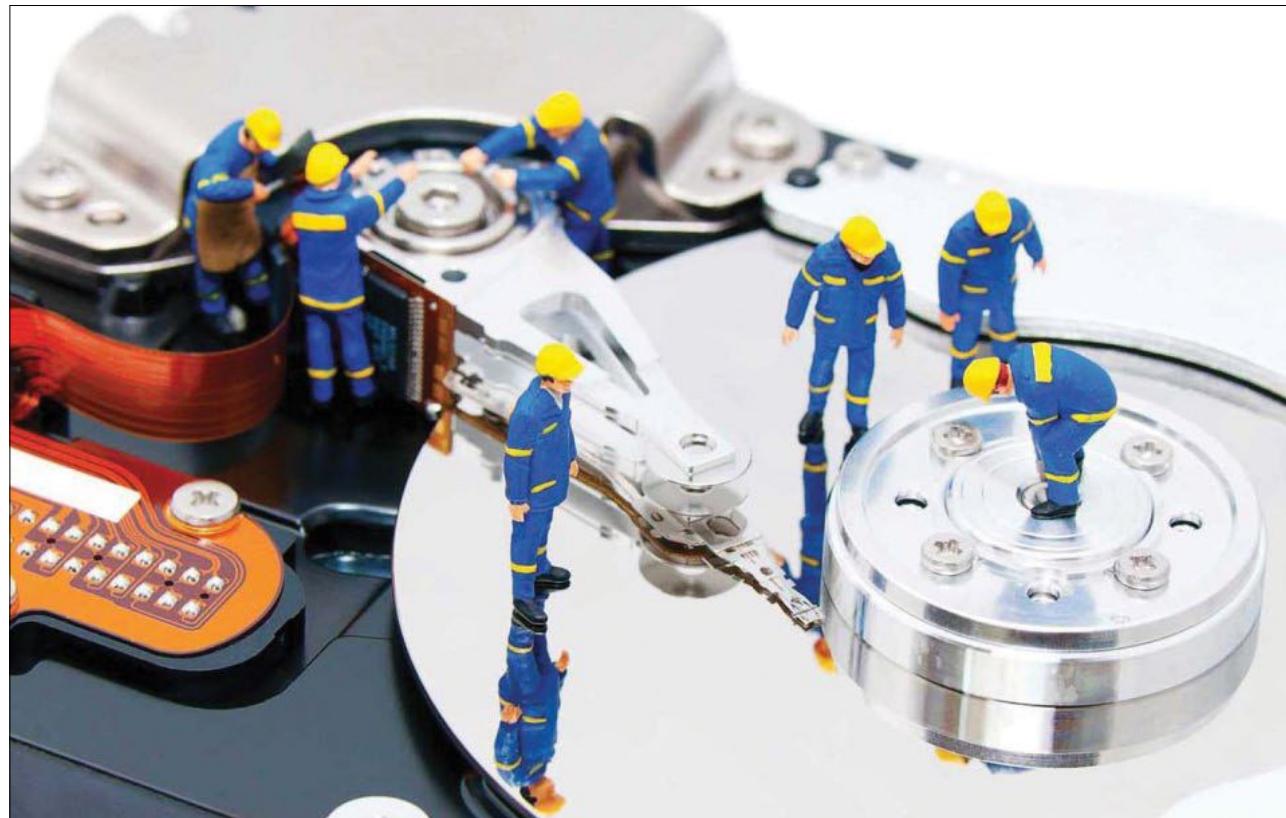


By: Ashish Sinha

The author is a software engineer based in Bengaluru. A software enthusiast by heart, he is passionate about using open source technology and sharing it with the world. He can be reached at ashi.sinha.87@gmail.com

An Introduction to Distributed Replicated Block Device

Distributed Replicated Block Device (DRBD) mirrors block devices among multiple hosts. The replication is visible to other applications on the host systems. Block devices such as hard disks, partitions, logical volumes, etc, can be mirrored to form high availability (HA) clusters.



In this article, let's explore some of the technology that works behind the cloud. To set up enterprise level cloud computing infrastructure, we not only need server grade hardware and cloud software but also something more. To ensure continuous availability and integrity of user data and application resources, we need a storage solution which can protect service continuity from individual server failures, network failures or even natural disasters. DRBD or Distributed Replicated Block Device is an open source solution that provides a highly available, efficient, distributed storage system for the cluster environment. It is developed and maintained by LINBIT, Austria and is available under GPLv2.

Architecture

Let's understand the DRBD architecture with the help of

an illustration (Figure 1) taken directly from the DRBD documentation. The two boxes represent two DRBD nodes forming a high availability (HA) cluster. A typical Linux distribution contains a stack of modules which more or less include a file system, a buffer cache, a disk scheduler and a disk driver to handle storage devices. The DRBD module sits in the middle of this stack, i.e., between the buffer cache and the disk scheduler.

The black arrow represents a typical flow of data in response to the normal application disk usage. The orange arrow represents data flow initiated by DRBD itself for replication purposes, from one node to another via the network. For the purposes of our discussion, let's assume the left node to be the primary and the right node to be the secondary DRBD node. Whenever a new disk block

is updated in the primary module, DRBD sends back the same block content to the secondary DRBD module in the peer node to replicate in its local disk, usually via a dedicated NIC. This way, DRBD keeps the data in both the nodes in sync. The primary node holds the service IP, i.e., the IP with which external applications communicate with the services provided. If the primary node fails, the secondary node takes over as the primary and the service IP by means of the HA service.

Terminology

Before we proceed further let's understand a few DRBD terms.

Resource: A resource is a set of things that form a replicated data set. It consists of a resource name, volumes, DRBD devices and a connection.

Resource name: Any arbitrary ASCII string excluding white space.

Volumes: Every resource consists of a set of volumes that share a single replication stream. Volumes contain the data set and some meta-data used by DRBD itself.

DRBD device: A virtual block device `/dev/drbdX`, corresponds to each volume in the resource, where `X` is the device's minor number.

Connection: A communication link between two nodes. Each resource will have exactly one connection.

Feature list

Though DRBD has matured into a system with a rich set of functionalities, let's look at some of the most important ones.

- Supports shared-secret authentication
- LVM (Logical Volume Manager) compatible
- Supports heartbeat/pacemaker resource agent integration
- Supports load balancing of read requests
- Automatic detection of the most up-to-date data after complete failure
- Delta resynchronisation
- Existing deployment can be configured with DRBD without losing data
- Automatic bandwidth management

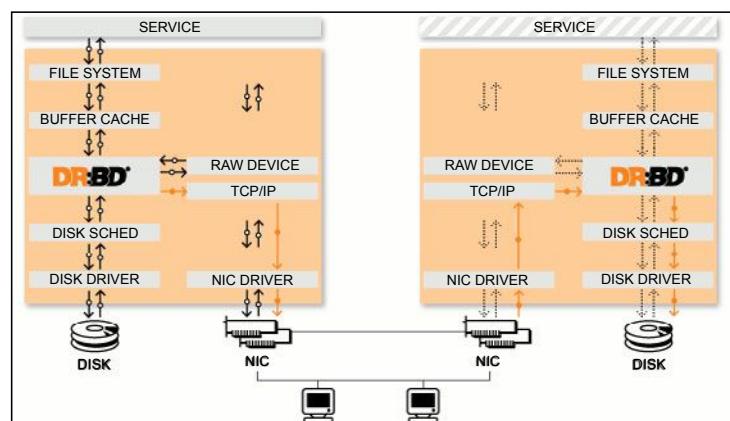


Figure 1: Position of DRBD within Linux

- Customisable tuning parameters
- Online data verification with peer

Components

DRBD functionalities are implemented in the following components:

drbd: This is a kernel module—the core—which actually does the work for us.

drbdadm: This is a userland administrative tool. It reads the configuration file and invokes the following tools with the required parameters, to get the job done.

drbdsetup: This is a userland administrative tool.

It configures the `drbd` kernel module. Usually, direct invocation by the user is not required. `drbdadmin` invokes whenever required.

drbdmeta: This userland administrative tool manipulates the meta data structure. Usually, direct invocation is not required.

Redundancy modes

DRBD can be set up in the following redundancy modes:

Single primary mode: In this mode, a DRBD resource takes the *primary* role on only one cluster member at a time. Since only the *primary* manipulates data, non-sharing file systems like ext3 and ext4 can be used in this configuration. This approach is used to set up a fail-over (HA) capable cluster.

Dual primary mode: In this mode, a DRBD resource takes the *primary* role on both the cluster members at a time. Since both nodes manipulate data, a sharing file system with distributed locking mechanisms like OCFS2 and GFC is required. This approach is used for load-balancing a cluster.

Replication modes

Different replication modes of DRBD are as follows:

Asynchronous (Protocol A): In this mode, a write operation on the *primary* is considered complete as soon as it is written on the local disk and is put in the TCP send buffer for replicating in peers. Data loss in the *secondary* may occur if the *secondary* fails before this write actually gets a chance to reach the *secondary* node disk. This is most often used in long distance replication set-ups where communication delay is a considerable factor.

Semi-synchronous (Protocol B): In this mode, a write operation on the *primary* is considered complete as soon as it is written on the local disk and the former has reached the peer node (not written to peer disk yet). Forced failure may not cause data loss, because the *primary* can detect that data has not reached a peer, so marks it as a failure and takes necessary action. But if the *secondary* fails after the data has reached it and before it is written to its local

disk, data loss will occur.

Synchronous (Protocol C): A write operation on the *primary* is considered complete only after both *primary* and *secondary* disks confirm the write. So, data protection is ensured in case of the previous two failure scenarios. This is the most common replication mode.

Installation and configuration

Now let's start our DRBD installation. For this set-up, I have taken two nodes running Ubuntu 14.04 LTS Server OS. The host IDs of the nodes are *drbd-server1* and *drbd-server2*. We also assume both the nodes have two NICs – the first is *eth0* for the *Service IP*, which will be accessed by external usage and managed by the HA cluster agent. The second NIC interface, i.e., *eth1* is used for DRBD replication purposes. Configured IP addresses are *10.0.6.100* and *10.0.6.101*, respectively. Configuring the HA agent is beyond the scope of this article.

Step 1: Edit */etc/network/interfaces* on *drbd-server1* and add the following:

```
auto eth1
iface eth1 inet static
address 10.0.6.100
netmask 255.255.255.0
network 10.0.6.0
broadcast 10.0.6.255
```

On *drbd-server2*, the configuration is similar except for its own IP *10.0.6.101*.

Step 2: Edit */etc/hosts* on both nodes and add the following:

```
10.0.6.100 drbd-server1
10.0.6.101 drbd-server2
```

Reboot both the nodes.

Step 3: Install the DRBD package on both the nodes, as follows:

```
sudo apt-get install drbd8-utils
```

Step 4: Check */etc/drbd.conf* for the following two entries, on both the nodes:

```
include "/etc/drbd.d/global_common.conf";
include "/etc/drbd.d/*.res";
```

Step 5: Edit */etc/drbd.d/global_common.conf* on both the nodes as follows:

```
global {
  usage-count yes;
```

www.opensourceforu.com

An **EFY GROUP**
PUBLICATION

Open Source Is Hot In The IT World



THE
COMPLETE
MAGAZINE
ON OPEN
SOURCE



➤ To find dealers to buy copy from news-stand,
visit: www.ezine.lfymag.com

➤ To subscribe, visit: pay.efyindia.com

➤ To buy an ezine edition,
visit: www.magzter.com choose Open Source For You

```

}
common {
net {
  protocol C;
}
}

```

The *usage-count yes* option enables collecting real-time statistics of DRBD, and *Protocol C* sets up a *synchronous mode* replication as described earlier in the *Replication Modes* section.

Step 6: Edit */etc/drbd.d/r0.res* on both the nodes:

```

resource r0 {
  device   /dev/drbd1;
  disk     /dev/sdb1;
  meta-disk /dev/sda2;
  on drbd-server1 {
    address  10.0.6.100:7789;
  }
  on drbd-server2 {
    address  10.0.6.101:7789;
  }
}

```

r0 is our resource name. Here, */dev/drbd1* is the virtual block device to be created by DRBD. The partition */dev/sdb1* holds our replicated data set. The partition */dev/sda2* holds meta information for internal use by DRBD. The port number 7789 is used by DRBD for replication connection, so make sure that it is not in use by any other application running on your system.

Step 7: Clear the meta-disk */dev/sda2* to avoid DRBD configuration issues:

```
sudo dd if=/dev/zero of=/dev/sda2
```

This takes some time, depending on the partition size.

Step 8: Create meta-data for the resource *r0*, as follows:

```
sudo drbdadm create-md r0
```

Step 9: Enable the resource *r0* as follows:

```
sudo drbdadm up r0
```

Step 10: Check the DRBD status in the */proc/drbd* virtual file system, as follows:

```
cat /proc/drbd
```

Check the disk state or *ds* field for ‘Inconsistent/Inconsistent’, which is normal at this stage since we have not done the initial synchronisation.

Step 11: Start the initial synchronisation from drbd-server1, to make it *primary* and other node *secondary*.

```
sudo drbdadm primary --force r0
```

Now, check the status in */proc/drbd* for the ongoing synchronisation. This step takes some time, depending on the disk size. On completion, the disk state *ds* shows ‘UpToDate/UpToDate’ and roles *ro* shows ‘Primary/Secondary’. Here, *primary* indicates that the local node resource is in *primary* mode and the peer is in *secondary* mode. Now, our DRBD set-up is ready for replication.

To test the set-up, let us mount our virtual block device */dev/drbd1* and create a file in it.

Let’s assume that the user name and group name of our system are the same, which is *drbd*.

```

mkdir $HOME/safedisk
sudo mount /dev/drbd1 $HOME/safedisk
sudo chgrp drbd $HOME/safedisk
sudo chown drbd $HOME/safedisk
cd $HOME/safedisk
echo "DRBD is up and running!" > myfile

```

Now shut down the *primary* to simulate a failure condition. Go to *secondary* (drbd-server2) and run the following command to change it to *primary*:

```
sudo drbdadm primary r0
```

Next, on *drbd-server2*, mount */dev/drbd1* to a local directory as described above and check *myfile* and its content. It should be the same as what we have written from the previous *primary* node.

This is only an introductory article on DRBD. Interested readers are advised to explore further to find out other features, configuration options and use-cases of DRBD. Online training is also available at <http://www.linbit.com/en/services/training> by LINBIT. 

References

- [1] <http://drbd.linbit.com/>
- [2] https://en.wikipedia.org/wiki/Distributed_Replicated_Block_Device
- [3] <https://help.ubuntu.com/lts/serverguide/drbd.html>

By: Ramaprasad Chakraborty

The author works at ISRO, Sriharikota, as a scientist engineer SD and has six years of industry experience. He is interested in real-time operating systems, controller redundancy mechanisms, networking, cyber forensics, database management systems and task automation with different levels of scripting and various open source technologies.

Open Source Frameworks for Big Data Processing: An Overview

This article discusses the Big Data processing ecosystem and the associated architectural stack. It investigates different frameworks suiting the various processing requirements of Big Data. It also delves into the frameworks at various layers of the stack such as storage, resource management, data processing, querying and machine learning.



The initial growth of data on the Internet was primarily driven by a greater population gaining access to the Web. This improved access was fuelled by the advent of a newer range of devices like smartphones and tablets. Riding on the first generation of data growth, a second wave of scaled up data production was unleashed mainly by the social media platforms that drove the growth upwards, exponentially. The collaborative nature of the information-sharing platforms contributed to a viral growth of the data shared via these platforms. The third wave of data generation is largely being led by the proliferation of intelligent connected devices and will lead to a scale of data generation that is unprecedented.

In addition, the advancements in scientific fields coupled with the availability of cheaper computing has led to newer applications in the fields of medical sciences, physics, astronomy, genetics, etc, where large volumes of data are collected and processed to validate hypotheses, and enable discoveries and inventions.

The huge growth in data acquisition and storage has led us to the next logical phase, which is to process that data to make sense of it. The need to process this vast volume of data has led to the demand for scalable and parallel systems that process the data at speed and scale. Open source technologies are a natural choice for the high performance computing needed for large scale data processing.

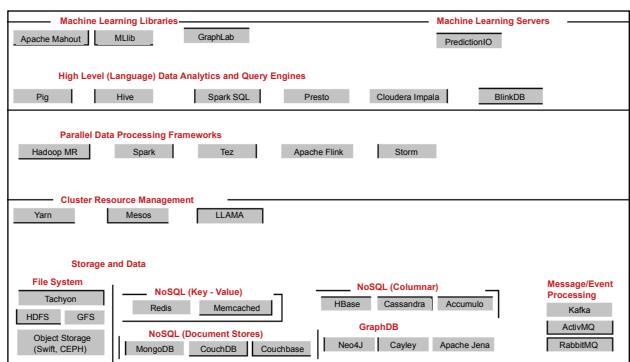


Figure 1: Component stack for Big Data processing

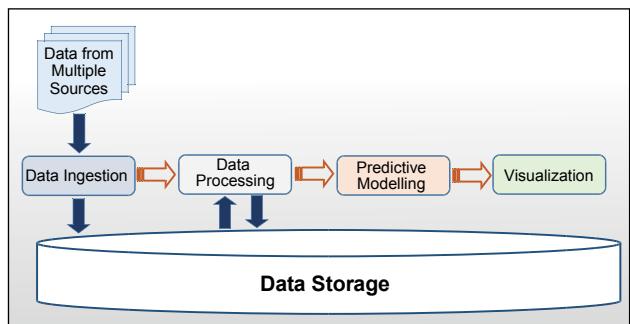


Figure 2: Big Data processing pipeline

This article aims to provide an overview of the frameworks and components available in open source, across different layers of a Big Data processing stack.

Component architecture stack for Big Data processing

As more and more Big Data, characterised by the three Vs - volume, velocity and variety - began to be generated and acquired, different systems started to evolve to tap the vast and diverse potential of this data. Although some of the systems converged in terms of the features they offered, they were all driven by different underlying design philosophies and, therefore, offered different alternatives. However, one of the guiding principles to develop enterprise data strategies would be to have a generic data storage layer as a data lake, which would allow different computing frameworks to work on the data stored for different processing use cases, and have the data shared across frameworks. Figure 1 illustrates a representational architectural stack for Big Data processing.

The stack can also be visualised as a pipeline consisting of multiple stages through which the data is driven, as can be seen in Figure 2. The unstructured and often schemaless raw data that is sourced from multiple sources, such as transactions, weblogs, open social sites, other linked data sources and databases, devices and instruments, could be in varying formats – such as textual data, images, video, audio,

etc. This data is then cleaned and often checked for any errors when ingested into the data storage layer. It is then processed, a task that might occur iteratively and interactively using the frameworks described in later sections. The processing could itself have multiple sub-stages and could revert data back into the storage layer at every iteration. It could be further explored and modelled using statistical algorithms to derive and validate hypotheses. These algorithms are trained on the data to learn the models, which can then be used for predictive modelling. The models could then be trained periodically as newer datasets flow into the system. The datasets are further used for exploratory analytics to discover unseen intelligence and insights. During the processing and exploratory processes, the processed datasets are visualised using visualisation tools to aid data understanding and for communicating to stakeholders.

This data in the storage layer could be reused by different stakeholders within an organisation. Big Data is typically undefined and most frameworks, as we will see later, have adapted to this aspect of it. In fact, this very feature is instrumental in the success of a framework.

Let us discuss some of the frameworks and libraries across these different layers.

The storage and data layer

Let's start with the storage and data layer, which is the most critical and the foundation of a Big Data stack. Big Data is typically characterised by its volume, requiring huge and conceptually unlimited storage capacities. Advances in technology, contributing to cheaper storage and compute resources, have resulted in the emergence of cluster storage and compute platforms. The platforms have unlocked the storage limitations and virtually enabled unlimited amounts of data storage. These platforms are not limited by the traditional paradigms of data modelling and schema designs. They are generally schema-free and allow the storage of all forms of data (structured, semi-structured and unstructured). This enables the creation of systems that are more dynamic and which enable analysts to explore the data without being limited by preconceived models. In this section, we will look at some of the popular cluster storage frameworks for Big Data.

HDFS (<https://hadoop.apache.org/>): This is a scalable, fault-tolerant distributed file system in the Hadoop ecosystem. HDFS is scaled by adding commodity servers into the clusters. The largest cluster size is known to be about 4500 nodes in a cluster with up to 128 petabytes of data. HDFS supports parallel reading and writing of data. The bandwidth in an HDFS system scales linearly with the number of nodes. There is built-in redundancy with multiple copies of data stored in the system. The files are broken into blocks and stored as files across the cluster. They are replicated for reliability.

HDFS has a master/slave architecture, wherein a cluster has a single component called NameNode, which acts as the master server. NameNode manages the file system namespace (files, directories and blocks, as well as their relationships). The namespace is stored in the memory and changes are persisted into the disk on a periodic basis.

In addition to that, there are slave components called DataNodes, usually one per node in the cluster. These processes manage the storage attached to the specific compute node they run on.

NoSQL databases (<http://nosql-database.org/>): As the Web has grown and become more accessible, it has become apparent that the existing relational database technologies are not equipped to handle the huge volumes and concurrency requirements of Web 2.0. To meet this need ‘Not only SQL’ databases have emerged as alternate data storage and management systems. While HDFS and Hadoop are data processing engines for analytics use cases, either in batch or real-time, NoSQL databases essentially serve as data storage layers for Web based front-end systems that need large concurrent data handling capacities.

Some of the key features that characterise these databases are that they are usually schema-free (or have a minimal schema structure), horizontally scalable, and rely on eventual consistency models rather than immediate consistency models.

There are four basic architectures of NoSQL databases that have emerged. These are:

- Key-Value stores are based on the data model of hash data structures or associative arrays. These systems were based on Amazon’s DynamoDB paper (<http://www.allthingsdistributed.com/files/amazon-dynamo-sosp2007.pdf>).
- Columnar databases (DBs) are based on Google’s BigTable paper (<http://research.google.com/archive/bigtable OSDI06.pdf>).
- The data model here is that each row can have its own schema, e.g., HBase and Cassandra.
- Document DBs are systems where the data is stored as documents, and each document is a collection of Key-Value pairs. Generally, these are expressed as JSON documents (e.g., MongoDB and CouchDB).
- Graph DBs are systems where the data models are represented as nodes, or as relationships between the nodes. Both nodes and relationships are represented in key-value pairs (e.g., Neo4J).

Tachyon (<http://tachyon-project.org/>): This is a platform that provides reliable in-memory data sharing across cluster frameworks and jobs. Tachyon essentially sits on top of storage platforms such as HDFS, thereby providing memory-centric data processing capabilities across cluster frameworks and jobs. Although some of the existing cluster computing frameworks like Spark, etc, have leveraged in-memory data

processing, there were three key shortcomings that motivated the development of Tachyon:

- Although jobs processed data within memory, sharing of data across jobs and frameworks was not achieved as the data was only available within the JVM context of the job.
- As the execution engine and storage was within the same JVM context, any execution engine crash led to the loss of data and necessitated re-computation.
- Data was replicated in-memory across jobs in certain cases leading to a larger data footprint, and heavier garbage collection.

Tachyon was developed to solve the above problems and was driven by a need to push lineage down to the storage layer. It enables storing of only one copy of data in the memory, which is made available across all frameworks such as Spark, MapReduce, etc. Moreover, fault tolerance was enabled by leveraging re-computations using lineage.

Data processing frameworks

Once the data is persisted into a storage layer, the next step is to process this data to derive insights. There are several frameworks that we will look at here.

The Apache Hadoop stack (<https://hadoop.apache.org/>) is the grand-daddy of Big Data processing frameworks, and has become the de-facto platform into which the technology has largely converged. The cost-effectiveness and the scalability of the platform is a perfect match for the needs of the large scale data processing in the industry. In addition, the reliability and the community support around the platform and the ecosystem have led to a wider adoption of the platform.

The Hadoop ecosystem has three main goals:

Scalability - enabling scaling to cater to larger requirements with just the addition of nodes to the cluster. This is further amplified by the fact that the framework relies on the local computation model to derive benefits from the simplified scalability model.

Flexibility – to provide the flexibility to store differently structured data formats. This is enabled by the ‘Schema on Read’ approach, which enables the system to store anything, and only decipher the schema at the time of reading the data, which is when it is required to know the data.

Efficiency – to ensure the cluster resources are optimally utilised for higher efficiency.

Hadoop MapReduce (<https://hadoop.apache.org/>) is the implementation of the *MapReduce* programming paradigm (popularised by a Google paper). This programming paradigm is modelled to process very large datasets in parallel, on large clusters, while ensuring reliability and fault-tolerance.

The *MapReduce()* paradigm itself is founded on the concept of a distributed file system which ensures reliability and scalability. A *MapReduce()* program consists of two procedures – *Map()* and *Reduce()*. The *Map()* procedure processes the input dataset in parallel and emits a processed

output. As the *Map()* phase happens across a very large distributed dataset, spread across a huge cluster of nodes, it is subsequently run through a *Reduce()* phase which aggregates the sorted dataset, coming in from multiple map nodes. This framework, along with the underlying HDFS system, enables processing of very large datasets running into Petabytes, spread across thousands of nodes.

Apache Flink (<https://flink.apache.org/>) is a data processing system that combines the scalability and power of the Hadoop HDFS layer along with the declarations and optimisations that are the cornerstone of relational database systems. Flink provides a runtime system, which is an alternative to the Hadoop MapReduce framework.

Apache Tez (<https://tez.apache.org/>) is a distributed data processing engine that sits on top of Yarn (Hadoop 2.0 ecosystem). Tez models the data processing workflows as Distributed Acyclic Graphs (DAGs). With this distinctive feature, Tez allows developers to intuitively model their complex data processing jobs as a pipeline of tasks, while leveraging the underlying resource management capabilities of the Hadoop 2.0 ecosystem.

Apache Spark (<https://spark.apache.org/>) is a distributed execution engine for Big Data processing that provides efficient abstractions to process large datasets in memory. While MapReduce on Yarn provides abstraction for using a cluster's computational resources, it lacks efficiency for iterative algorithms and interactive data mining—algorithms that need to reuse data in between computations. Spark implements in-memory fault tolerant data abstractions in the form of RDDs (Resilient Distributed Datasets), which are parallel data structures stored in memory. RDDs provide fault-tolerance by tracking transformations (lineage) rather than changing actual data. In case a partition has to be recovered after loss, the transformations need to be applied on just that dataset. This is far more efficient than replicating datasets across nodes for fault tolerance, and this is supposedly 100x faster than Hadoop MR.

Spark also provides a unified framework for batch processing, stream data processing, interactive data mining and includes APIs in Java, Scala and Python. It provides an interactive shell for faster querying capabilities, libraries for machine learning (MLlib and GraphX), an API for graph data processing, SparkSQL (a declarative query language), and SparkStreaming (a streaming API for stream data processing).

SparkStreaming is a system for processing event streams in real-time. SparkStreaming treats streaming as processing of datasets in micro-batches. The incoming stream is divided into batches of configured number of seconds. These batches are fed into the underlying Spark system and are processed the same way as in the Spark batch programming paradigm. This makes it possible to achieve the very low latencies needed for stream

processing, and at the same time integrate batch processing with real-time stream processing.

Apache Storm (<https://storm.apache.org/>) is a system for processing continuous streams of data in real-time. It is highly scalable, fault tolerant, and ensures the notion of guaranteed processing so that no events are lost. While Hadoop provides the framework for batch processing of data, Storm does the same for streaming event data.

It provides Directed Acyclic Graph (DAG) processing for defining the data processing pipeline or topology using a notion of spouts (input data sources) and bolts. Streams are tuples that flow through these processing pipelines.

A Storm cluster consists of three components:

- **Nimbus**, which runs on the master node and is responsible for distribution of work amongst the worker processes.
- **Supervisor** daemons run on the worker nodes, listen to the tasks assigned, and manage the worker processes to start/stop them as needed, to get the work done.
- **Zookeeper** handles the co-ordination between Nimbus and Supervisors, and maintains the state for fault-tolerance.

Higher level languages for analytics and querying

As cluster programming frameworks evolved to solve the Big Data processing problems, another problem started to emerge as more and more real life use cases were attempted. Programming using these computing frameworks got increasingly complex, and became difficult to maintain. The skill scalability became another matter of concern, as there were a lot of people available with domain expertise familiar with skills such as SQL and scripting. As a result, higher level programming abstractions for the cluster computing frameworks began to emerge, that abstracted the low level programming APIs. Some of these frameworks are discussed in this section.

Hive (<https://hive.apache.org/>) and **Pig** (<https://pig.apache.org/>) are higher level language implementations for MapReduce. The language interface internally generates MapReduce programs from the queries written in the high level languages, thereby abstracting the underlying nitty-gritty of MapReduce and HDFS.

While Pig implements PigLatin, which is a procedural-like language interface, Hive provides the Hive Query Language (HQL), which is a declarative and SQL-like language interface.

Pig lends itself well to writing data processing pipelines for iterative processing scenarios. Hive, with a declarative SQL-like language, is more usable for ad hoc data querying, explorative analytics and BI.

BlinkDB (<http://blinkdb.org/>) is a recent entrant into the Big Data processing ecosystem. It provides a platform for interactive query processing that supports approximate

queries on Big Data. As data volumes have been growing exponentially, there has been an increasing amount of continuing research happening in this space to create computing models that reduce latency. Apache Spark was an effort in that direction, which worked on reducing latency using in-memory data structures.

Blink DB went further to squeeze the latency benchmarks by driving a notion of approximate queries. There were several industry use cases that appeared to tolerate some error in the answers, provided it was faster. BlinkDB does this by running queries against samples of original datasets rather than the entire datasets. The framework makes it possible to define the acceptable error bounds for queries, or specify a time constraint. The system processes the query based on these constraints and returns results within the given bounds.

BlinkDB leverages the notion of the statistical property – ‘sampling error’ – which does not vary with the population size, but rather depends on the sample size. So the same sample size should hold reasonably well with increasing data sizes. This insight leads to an incredible improvement in performance. As the time taken in query processing is mostly I/O bound, the processing time can be increased by as much as 10x with a sample size of 10 per cent of the original data, with an error of less than 0.02 per cent.

BlinkDB is built on the Hive Query engine, and supports both Hadoop MR as well as Apache Shark execution engines. BlinkDB provides an interface that abstracts the complexities of approximation, and provides an SQL-like construct with support for standard aggregates, filters and groupBy, joins and nested queries, apart from user-defined functions with machine language primitives.

Cluster resource management frameworks

Cluster resource management is one of the key components in the Big Data processing stack. Successful frameworks that have emerged have been able to combine generality in supporting different frameworks with disparate processing requirements as well as the robustness to handle volume and recovery seamlessly. A generic framework will avoid the need to replicate massive amounts of data between different disparate frameworks within a cluster. It is also very important to provide interfaces at this layer that enable ease of administration and use. We will discuss a couple of frameworks that have shown promise at this layer.

Apache Hadoop Yarn (<https://hadoop.apache.org/>): Hadoop 1.0 was written solely as an engine for the MapReduce paradigm. As Hadoop became widely accepted as a platform for distributed Big Data batch processing systems, requirements grew for other computing patterns like message passing interfaces, graph processing, real-time stream processing, ad hoc and iterative processing, etc. MapReduce, as a programming pattern, did not support these kinds of requirements, and newer (as well as other existing) frameworks

www.indiaelectronicsweek.com

An **EFY GROUP** EVENT

SHOWCASING LATEST T&M PRODUCTS & LEADING T&M BRANDS

Attend India's First LIVE Demonstration of Latest T&M Equipment



TEST & Measurement India

January 11th-13th, 2016 • Bengaluru

To register, visit: register.efyexpo.com or call 011-26810601/02/03

To increase your sales by exhibiting in this event,
Call: Jyotsna (011-40596605) or write to us at: efyexpo@efy.in

Co-located at

INDIA ELECTRONICS WEEK

January 11th-13th, 2016
Bengaluru

started to evolve. Also, HDFS was widely accepted for Big Data storage. It did not make sense to replicate data for other frameworks. The Hadoop community worked on re-hauling the platform to take it beyond MapReduce. The result of this was Hadoop 2.0, which separated resource management from application management. The resource management system was named Yarn.

Yarn is again a master-slave architecture, with the resource manager acting as a master that manages the resource assignments to the different applications on the cluster. The slave component is called the NodeManager, runs on every node in the cluster, and is responsible for launching the compute containers needed by the application.

The ApplicationMaster is the framework-specific entity. It is responsible for negotiating resources from the ResourceManager and working with the node manager to submit and monitor the application tasks.

This decoupling allowed other frameworks to work alongside MapReduce, accessing and sharing data on the same cluster, thereby helping to improve cluster utilisation.

Apache Mesos (<http://mesos.apache.org/>) is a generic cluster resource management framework that can manage every resource in the data centre. Mesos differs from Yarn in the way the scheduling works. Mesos implements a two-level scheduling mechanism, where the master makes resource offers to the framework schedulers, and the frameworks decide whether to accept or decline them. This model enables Mesos to become very scalable and generic, and allows frameworks to meet specific goals such as data locality really well.

Mesos is a master/slave architecture with the Mesos master running on one of the nodes, and is shadowed by several standby masters that can takeover in case of a failure. The master manages the slave processes on the cluster nodes and the frameworks that run tasks on the nodes. The framework running on Mesos has two components: a scheduler that registers with the master, and the framework executor that launches on the Mesos slave nodes. In Mesos, the slave nodes report to the master about the available resources on offer. The Mesos master looks up the allocation policies and offers the resources to the framework as per the policy. The framework, based on its goal and the tasks that need to be run, accepts the offer completely, partially or can even decline it. It sends back a response with the acceptance and the tasks to be run, if any. The Mesos master forwards the tasks to the corresponding slaves, which allocate the offered resources to the executor, and the executor in turn launches the tasks.

Machine learning libraries

Big Data would not be worth the effort if it didn't

provide a business value at the end. Machine learning programs the systems to learn or process large amounts of data, and be able to apply the learnings to predict outcomes on an unseen input dataset. Machine learning systems have enabled several real world use cases such as targeted ad campaigns, recommendation engines, 'next best offer/action' scenarios, self-learning autonomous systems, etc. We will look at a few of the frameworks in this space.

Apache Mahout (<http://mahout.apache.org/>) aims to provide a scalable machine learning platform with the implementation of several algorithms out-of-the-box, and provides a framework for implementing custom algorithms as well. Although, Apache Mahout was one of the earliest ML libraries, it was originally written for the MapReduce programming paradigm. However, MapReduce was not very well suited for the iterative nature of machine learning algorithms and hence did not find great success. However, after Spark started gaining momentum, Mahout has been ported to Apache Spark, rebranded as Spark MLLib, and has been discontinued on Hadoop MapReduce.

Spark MLLib (<https://spark.apache.org/mllib/>) is a scalable machine learning platform, which is written on top of Spark and is available as an extension of the Spark Core execution engine. Spark MLLib has an advantage as it has been implemented as a native extension to Spark Core. Spark MLLib has several algorithms written for ML problems such as classification, regression, collaborative filtering, clustering, decomposition, etc.

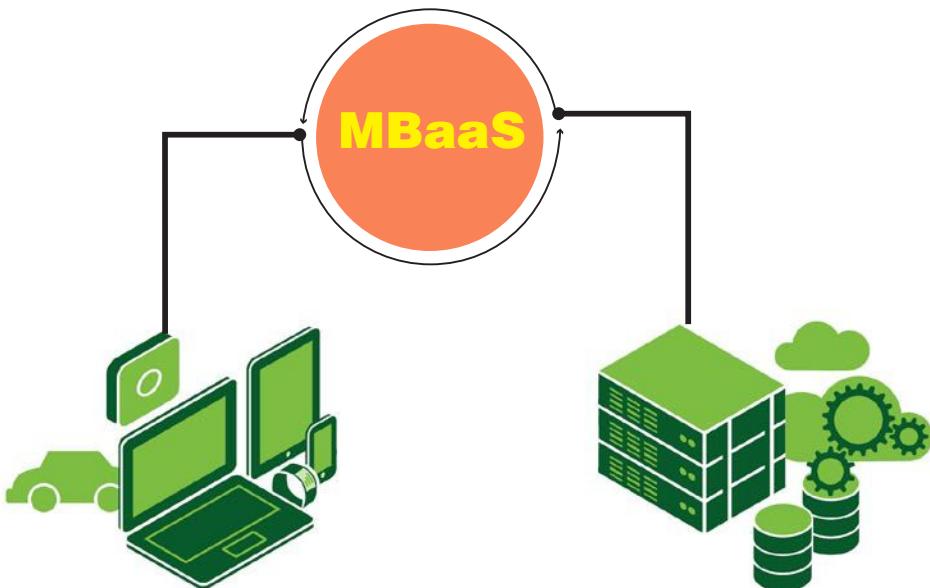
PredictionIO (<http://prediction.io/>) is a scalable machine learning server that provides a framework enabling faster prototyping and productionising of machine learning applications. It is built on Apache Spark and leverages Spark MLLib to provide implementation templates of several machine learning algorithms. It provides an interface to expose the trained prediction model as a service through an event server based architecture. It also provides a means to persist the trained models in a distributed environment. The events generated are collected in real-time and can be used to retrain the model as a batch job. The client application can query the service over REST APIs and get the predicted results back in a JSON response. 

By: Subhash Bylaiah and Dr B. Thangaraju

Subhash Bylaiah is a senior technologist with over 15 years of experience in database systems and application development. His current assignment is with Wipro Technologies. He can be reached at subhash.bylaiah@wipro.com.

Dr B. Thangaraju is an open source software (OSS) evangelist, who works at Talent Transformation, Wipro Technologies, Bangalore and he can be reached at balat.raju@wipro.com.

How MBaaS Providers are Simplifying App Development



Mobile app development requires a considerable amount of back-end services to sustain the app. Rather than expend energy and time on the complexities of back-end services, developers would prefer to devote their efforts to developing the apps. Cloud BaaS (Backend as a Service) or MBaaS (Mobile Backend as a Service) providers now come to the rescue.

When we think of software or of Web and mobile app development, two things that come to mind are the server side and the client side. To put it in simple terms, the client side is the front end or interface screen(s) that the end user or client will see and interact with, whereas the server side is broadly a combination of application business logic and data management. With the availability of cloud computing, many of the server side complexities are handled by cloud service providers, as shown in Figure 1.

Software development is now moving towards app development, where the client side development requires more focus due to the variety of devices and platforms in use. This has led to reducing the interaction with servers and a whole new cloud service model has come into being, i.e., the Backend as a Service (BaaS) or Mobile Backend as a Service (MBaaS). This is a unique model, which enables mobile and Web developers to use cloud infrastructure as the back-end, hence freeing them from the headache of back-end designing and integration. It is like infrastructure that can be turned on just for mobile app development. As shown in Figure 2, mobile app development primarily comprises four activities. The major challenges for developers are with back-end integration and most of

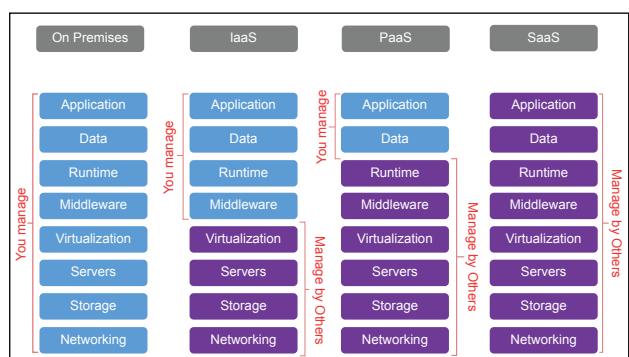


Figure 1: Separation of responsibilities

the effort is focused on this aspect, due to which application design and development suffer. The same applies for Web development also.

Creating a mobile application that works well is a complex and time-consuming task and needs a

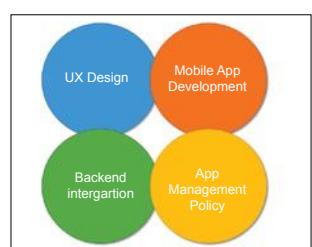


Figure 2: Typical tasks in mobile app development

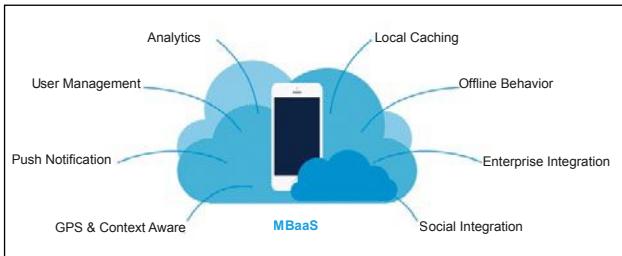


Figure 3: Features of MBaaS (source: <http://rapidvaluesolutions.com/whitepapers/How-MBaaS-is-Shaping-up-Enterprise-Mobility-Space.html>)

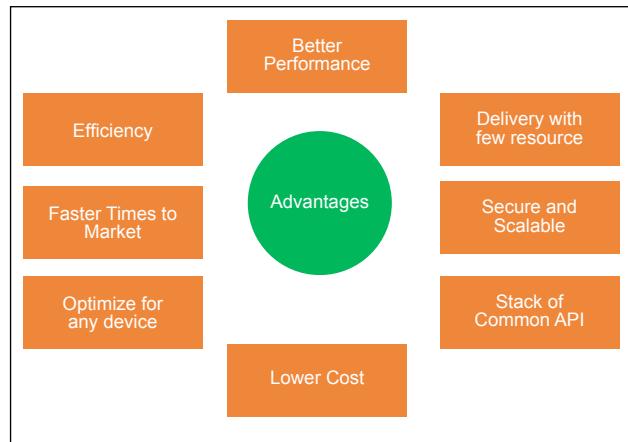


Figure 4: Advantages of MBaaS

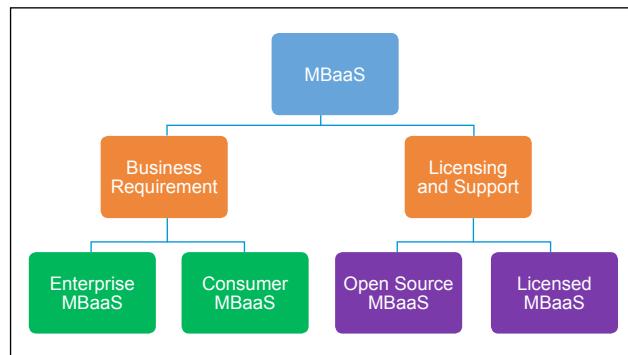


Figure 5: Categorisation of MBaaS

lot of planning to ensure proper execution. Some studies show that up to 80 per cent of the implementation effort is dedicated to back-end development.

MBaaS comes to the rescue as a server side technology for mobile apps, as it provides custom SDKs and APIs, as shown in Figure 3.

MBaaS is gaining popularity not only in the developer community but also in enterprises. According to Gartner, by 2016, 40 per cent of the mobile app development projects will leverage mobile back-end services. TechNavio's analysts also forecast that the global BaaS market, which was valued at US\$ 0.87 million in 2014 is expected to reach US\$ 29.16 billion by 2019, growing at a CAGR of 101.88 per cent.



Figure 6: Major players in MBaaS

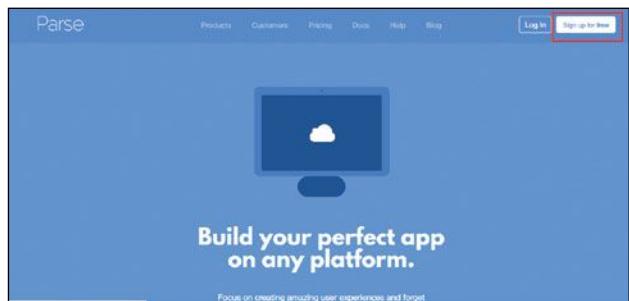


Figure 7: Home page of parse.com



Figure 8: Options for apps in parse.com



Figure 9: Selecting the environment

Why use it?

Mobile Backend as a Service (MBaaS) reduces the time and complexity involved in building mobile applications



Figure 10: Selecting a platform



Figure 11: Select a language

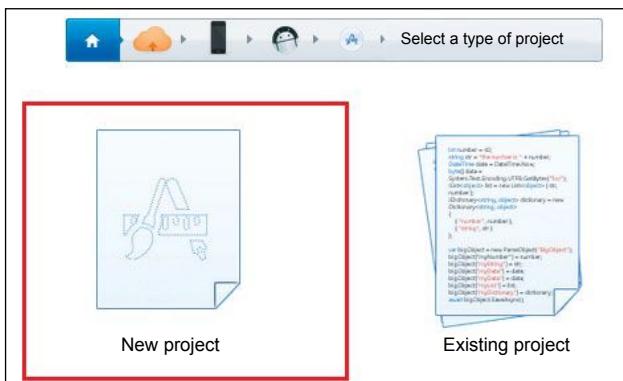


Figure 12: Selecting a type of project

by providing pre-built, efficient, cloud-hosted components for mobile back-ends. It allows developers to focus on core features.

Major players

MBaaS can be grouped in different ways according to business requirements or based on licensing and support models.

Some major MBaaS players are shown in Figure 6.

Apart from the few companies shown in Figure 6, there are many other MBaaS or BaaS providers in the market and their numbers are growing. It should be understood that no single service provider fits all sizes. An MBaaS provider is selected as per the specific and unique needs of individual organisations.

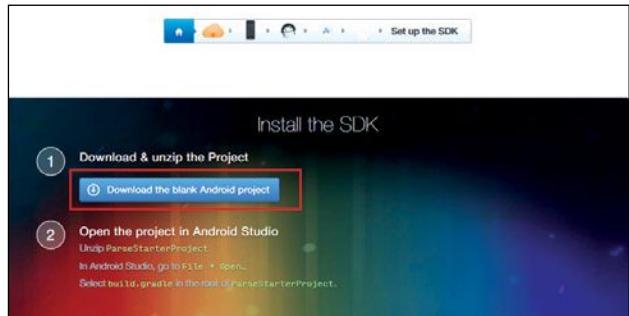


Figure 13 : SDK set-up



Figure 14: Code for testing

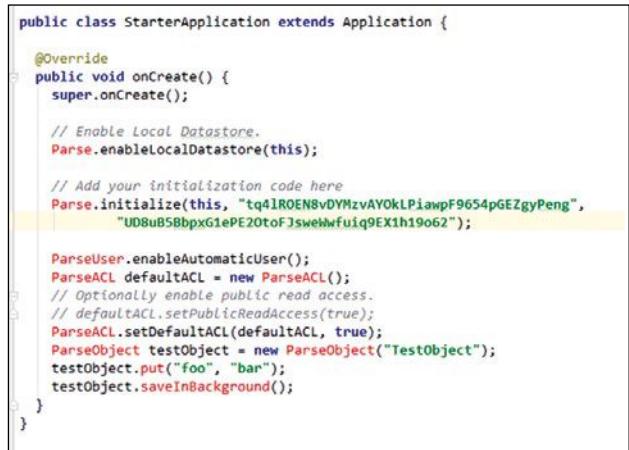


Figure 15: StarterApplication.java code

Some common features that are desirable in MBaaS are:

1. *User management APIs*, which save time when developing sign-ups, logins, email verification, etc
2. *Support for different authentication mechanisms*, which provides the social site's login using OAuth, etc
3. *Mobile app performance data*, which provide analytics and the dashboard for performance tuning
4. *Data storage considerations*, which refers to NoSQL or Relational or any kind of data storage
5. *Data analysis capabilities for evaluating user activity* to check push campaigns, user responses by different categories, etc



Workshops

Drupal 8
in a day

BigData
(what, why, how)

Cloud
in a day

Software Architecture:
Principles, Patterns, and Practices

Shifting to Functional Programming:
Lambdas for Java Developers

Cross Platform Mobile Application Development:
Consumer & Enterprise Applications

Tizen Web App Development -
A Hands On Session

Design, Deploy and Manage APIs using
WSO2 API Management Platform



TRACKS

Mobile App Dev Day

Cloud Day

FOSS for Everyone

WebApp Dev Day

IT Infrastructure Day

Kernel Day

Database Day

Open Source in IoT

OpenStack Mini Conference

IT Implementation Success Stories





Is the **Closed**
Code
Taming the
'Creative
Lion'
within you?

Come & Discuss **OPEN SOURCE**



12th Edition

OPEN
SOURCE INDIA

NIMHANS Convention Center

19 - 20
November
2015

BENGALURU

Asia's largest
conference on
Open Source

2 Days. 10 Tracks. 8+ Workshops. 40+ Speakers

For more details, call on 011-40596605 or email us at info@osidays.com

Platinum Partner

Gold Partner

Knowledge Partner

Developer
Engagement Partner

Associate Partners

'FOSS For
Everyone' Partner

Lanyard
Partner



edureka!





Figure 16: What happens after clicking on the *Test* button

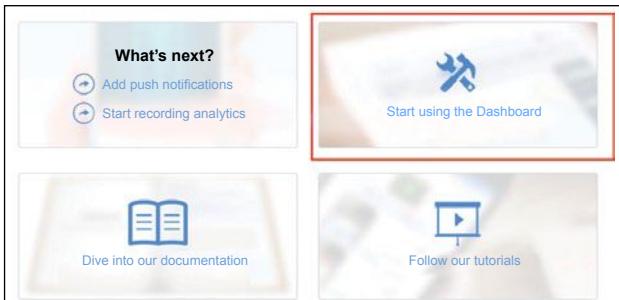


Figure 17: After testing the next step options

Pricing

The major factors for MBaaS pricing are:

- The number of API calls
- The type and size of storage
- Active user pricing
- The number of application builds
- Premium fees for analytics
- Feature based pricing – for push mail, chat, sync, email, etc
- Support and maintenance costs

Case study: Parse.com

Parse is a start-up company that was acquired by Facebook and provides MBaaS. According to a case study by Amazon, Parse provides server management for over 180,000 Android, iOS and Windows mobile applications, which run on more than 200 million mobile devices. One can see the hundreds of Parse customers at <https://parse.com/customers>.

Sample application

Using Parse is relatively easy. Open parse.com and click on *Sign up for free*, which is shown in Figure 7. Fill in the necessary details and create an application name of your choice. This demo will show a very simple and primitive application to demonstrate how easily data can be stored on the cloud using Parse.

Fill in the necessary details and click on *Signup*. Once registered, log in to parse.com, which will show what Parse is capable of in terms of a BaaS (Figure 8).

Select the data icon so that we can store some data. Parse will ask for environments (Figure 9).

Parse supports a variety of environments for which it can provide back-end as a service. Select *Mobile* as an environment, which will display the available platforms in the mobile environment as shown in Figure 10.

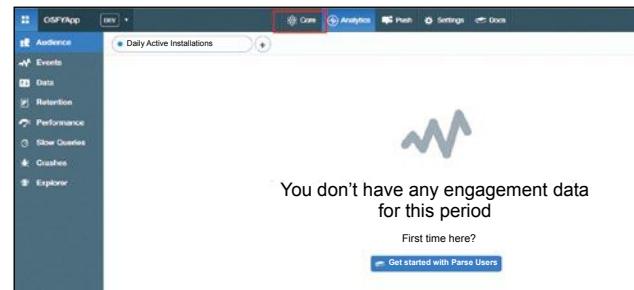


Figure 18: Dashboard view

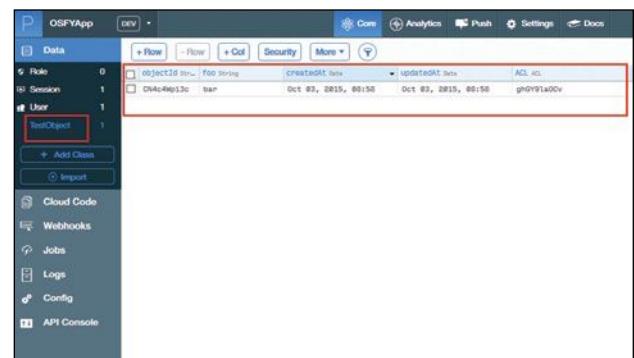


Figure 19: Data view

Select *Android* as the platform. It will display language options as shown in Figure 11.

Select *Native (Java)*, and you will be asked whether to start with a new or existing project as shown in Figure 12.

Select *New Project* which will show some options for the SDK as shown in Figure 13.

Download the blank project, and import the same in Android Studio.

At the bottom of the page, there is dummy code to store the data as shown in Figure 14.

Open Android Studio, find the *StarterApplication.java* file and paste the highlighted code of the *onCreate* method (Figure 15).

Run the application in the emulator which will store the code in the parse.com site. On the same page of parse.com there is a *Test* button and clicking on it will show the data creation (Figure 16).

To view the data later, click on the dashboard option as shown in Figure 17.

The dashboard will open in *Analytics view*. Click on *Core* as shown in Figure 18.

Core click on *TestObject* and the data that is saved from the Android app is visible (Figure 19). 

By: Ashish Singh Bhatia

The author is a technology enthusiast and a FOSS fan. He loves to explore new technology and work on Python and Java. He can be reached at ast.bhatia@gmail.com and blogs at <https://openfreeidea.wordpress.com/>

NFV and the Service Provider Cloud

Network Functions Virtualisation (NFV) involves virtualising entire network node functions into entities that may be connected or chained together, to create networking services.

This article examines the rise of Network Functions Virtualisation or NFV (which is built using cloud technologies in the service provider network), provides an overview of the ETSI reference architectural framework for NFV, and concludes with the role that open source projects play in this emerging area.

The service provider's dilemma

The demand for networks is increasing. Propelled by smartphones and social networking applications, people now produce and consume more pictures and videos every day resulting in an exponentially growing volume of traffic. With increasing network penetration, more and more people and devices access the

Internet every day. It is estimated that there are approximately 9 billion devices connected to the Internet already. While this may seem huge, technologists predict an even more rapid increase in Internet connected devices (to around 50 billion devices by 2020, according to some reports), a trend now called the Internet of Things (IoT).

While this appears to put network (Internet) service providers in a sweet spot, there is a problem. The revenues do not grow as fast, since more bandwidth and downloads need to be provided at lower costs. Handling the huge increase in scale (number of devices, number of flows) and throughput requires large capacity addition. They need to buy more networking equipment, wire it together, manage and operate it. Today's network design does not make this easy.

This results in a dilemma—not catering to the rising network

demand would result in losing market share, customer mind share and a leadership position; but to cater to the growth in demand would require huge investments in capacity, particularly for peak loads, and also involve additional management and operational complexities and costs.

Networks today

For the layman, the service provider network is an interconnection of routers and switches. In reality, diverse networking equipment is used in a service provider network. In addition to routers and switches, there are traffic load balancers, CG-NAT devices, content caching/serving devices (CDN), VPN boxes, a variety of security equipment starting with firewalls to intrusion detection and prevention devices to more complex threat management equipment, and many more. Figure 1 provides a conceptual illustration showcasing the diversity of equipment and their interconnections.

The service provider needs to buy the best-of-breed equipment from a variety of device vendors. Typically, each device is configured manually via Command Line Interface (CLI) by network experts trained for a particular category of devices. The way the devices are interconnected is static and limits the service flexibility that can be provided. Adding new devices or services into the mix is a daunting task.

Factors behind the rise of NFV

Server virtualisation: A little more than a decade ago, the server world underwent a transformation. It began when virtualisation technologies began to mature with support from CPU vendors. Instead of virtual machines (VMs) running on CPU emulators such as Bochs (<http://bochs.sourceforge.net/>), CPUs started to provide ways to run the VM's instructions natively. This boosted the performance of applications running in VMs to be nearly the same as when running natively. This enabled enterprises to consider server virtualisation as an option.

The migration from running applications on physical servers to running them within virtual machines enabled enterprises to make better use of their physical servers. By running services within VMs and orchestrating virtual machine creation/deletion and steering traffic to the new VM instances, enterprises were able to get elasticity, i.e., the ability to deal with fluctuating demands in an optimal manner. More VMs were provided to services seeing higher loads and less VMs to services with low load conditions.

Service providers too face similar problems. Load on a network is not the same at all times. Hence, they began to seriously consider leveraging virtualisation technologies to solve their network problems. If network services such as NAT or IDP/IPS could be run in VMs on server platforms, they realised they could save on hardware costs and reap the same benefits as enterprises.

I/O virtualisation: Once the ability to run a VM's CPU instructions at near native speeds was provided, CPU vendors started adding the ability to virtualise I/O. Until then, the I/O performance was constrained by the hypervisor's capacity or the host-OS's capacity. Thus, virtualisation was suited only for

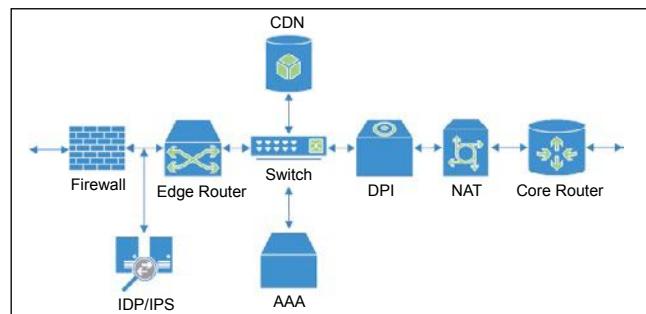


Figure 1: Today's service provider network

compute-heavy workloads. But, networking applications are I/O heavy.

I/O virtualisation technologies such as SR-IOV and the ability to DMA (direct memory access) from physical devices directly into a virtual machine's address space provides a way for VMs to access the physical Ethernet ports directly, rather than via the hypervisor's device emulation. This enables the VM based applications to utilise the full bandwidth of the network ports.

Software defined networking (SDN): SDN is an emerging trend to make networks programmable. SDN is about providing applications both visibility and the ability to control how traffic flows in the network. To that end, applications are provided the required abstractions of the network, information about traffic flowing via the network and the APIs to instruct how traffic must flow.

To provide a network level view, SDN involves the separation of the data plane and control plane, moving the control plane out of the network equipment, and centralising the control plane. For interoperability and vendor independence, the data plane and control plane communicate via a standards based protocol, the most popular one being OpenFlow. It abstracts the data plane as a set of match-action rules, where N-tuple from the Ethernet-frame is looked up and the action defined by the rule is performed.

In a traditional network, each device in the network maintains its own view of the network and exchanges messages with its peers via routing protocols, so that all devices maintain the same view of the network. Configuring the control plane protocols requires special knowledge and training. When new network services are to be introduced into the network, the control plane configuration needs to be tweaked and tested, in order to ensure that there is no unintended impact on the network. This requires planning, time and effort, and hence a cost for the service providers.

In SDN, the controller takes the role of the control plane. The job of the SDN controller is to obtain information from the SDN data plane elements, build a topology, and translate the information from the network into an abstract form that can be provided to applications. Similarly, instructions received from the applications are translated into actions for each piece of equipment in the SDN network.

Coming to cloud deployments, elasticity – the ability to create or remove VMs based on load - is a key requirement for operational efficiency. This requires that traffic from the data centre gateway be steered to the new VM instance, when a VM

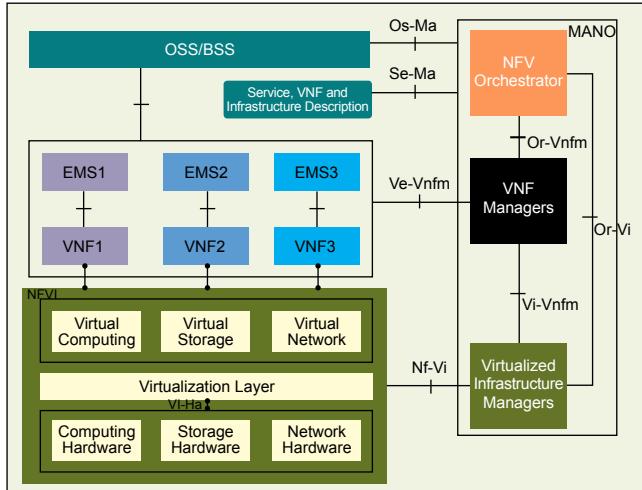


Figure 2: ETSI NFV architecture reference model

is created, or steered to other VM instances, when a VM is removed. SDN enables rapid traffic steering by helping the SDN controllers to modify the actions to be taken by SDN data plane elements for individual traffic flows.

The need for rich services

With the cost of bandwidth being lowered, service providers aim to improve revenues by providing a rich set of value added services. In current networking equipment, services are typically handled in a separate ‘service plane’, kept in the same chassis as the data and control plane components. With increasing line-card capacity, the service plane complex becomes more and more overwhelmed. To remedy this, services are provided via dedicated high performance equipment, as shown in Figure 1. These service equipment are chained together to provide a set of services, and this is called ‘service chaining’. Today, it is static in nature and depends on how the devices are connected.

However, once services are moved out of the chassis, it also creates the possibility to run the same services as virtual instances running in server platforms. With SDN, we can create service chains dynamically. This provides a lot of flexibility and operational convenience to service providers. Thus, ‘service chaining’ is a primary driver for service providers to consider NFV, while retaining the investments already made by them in high-capacity routers and switches.

Other factors

The service providers are looking for standards based solutions so that they can get the best-of-breed equipment and avoid vendor lock-in. As the network becomes software-centric rather than hardware-centric, they are considering open source solutions as well.

The emergence of Ethernet, once considered a LAN technology, into the WAN arena also helps. Ethernet speeds have been increasing with constant innovations in the physical (optical) layers. Today we talk of servers with 40G ports and

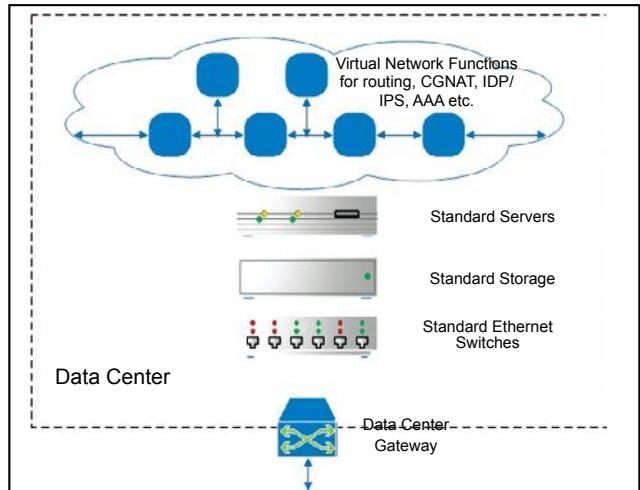


Figure 3: Service provider networks in the future (NFV based)

even 100G ports. This means that servers running virtualised network appliances no longer need other equipment to terminate protocols such as SONET and ATM. They can directly terminate network traffic and provide service.

ETSI architecture for NFV

In 2012, the European Telecommunications Standards Institute (ETSI) created an industry standards group (ISG), supported by some service providers, to define requirements, use-cases, architectural references and specifications in the area of NFV.

Since then, ETSI has published various documents, defining a reference architectural model, shown in Figure 2, and various recommendations to enable NFV adoption and interoperability.

The NFVI block shown at the bottom of Figure 2 comprises the hardware, i.e., physical servers, storage devices and physical network comprising networking hardware and interconnects. Above this lies the hypervisor layer, which provides virtualised compute, storage and networks for running applications. Various Layer 2 tunnelling technologies (e.g., VXLAN, NVGRE) help create a virtual network overlay over the physical network.

The virtual network functions (VNFs), managed by their respective element management systems (EMS), run above this virtualised infrastructure. The VNFs provide the functionality that is currently provided via physical network hardware.

The NFV management and orchestration layer (MANO) comprises the orchestrator, virtual machine manager and virtualised infrastructure manager. The orchestrator works with the OSS/BSS system to determine load and network services for which VNFs need to be created or deleted. It works with the virtual machine manager to create/remove VNFs and with a virtual infrastructure manager to assign hardware, storage and network resources to the VNFs. It also steers traffic to the VNFs and creates service chains by connecting multiple service VNFs.

Figure 3 illustrates what a future service provider network based on cloud technologies would look like.

Open source projects in NFV

Open source projects are playing a key role in the evolution and adoption of cloud technologies in general, and NFV in particular. These projects are also well supported by the industry.

In this section, we will look at some well-known open source projects relating to NFV. A brief about each project is provided along with a mention as to where they fit in the ETSI NFV reference architecture.

There is a rich open source ecosystem covering various aspects of virtualisation, SDN, NFV and the cloud. Thus, this section is not and cannot be exhaustive. Nor is it intended to be an endorsement or otherwise of any specific project mentioned or omitted here.

QEMU/KVM (wiki.qemu.org, www.linux-kvm.org):

KVM, together with QEMU, is a basic hypervisor providing both processor and board virtualisation. It is also integrated into the Linux kernel and hence is widely available and increasingly used in many cloud deployments.

In the ETSI reference architecture, QEMU/KVM provides the virtual compute capability.

Data Plane Development Kit - DPDK (www.dpdk.org): The data plane deals with handling packets/frames and applying the required functionality. The performance of networking equipment depends on the performance of the data plane. To build a high performance data plane, network equipment vendors have been using custom ASICs or network processors (NPUs). The NPUs are specialised processors optimised for packet processing.

However, with advances in CPU architecture and the advent of multi-core x86 processors, it is possible to use DPDK to create reasonably high-performance data planes running on x86 processors. DPDK provides a set of libraries, drivers and template applications to help create a data plane application.

In the ETSI reference architecture, DPDK is a building block to create VNFs.

Open vSwitch (openvswitch.org): The Open vSwitch project provides an open source virtual Ethernet switch. VNFs can comprise multiple VMs that use a private Ethernet network to interconnect them. This would be the case when a legacy chassis based network application is being virtualised.

Open vSwitch can also be used to connect multiple VNFs into a single virtual Ethernet switch, thereby providing the ability to locate VNFs in different physical nodes. It provides a virtual LAN overlay over the physical Ethernet network of a data centre.

In the ETSI reference architecture, Open vSwitch comes under the virtual network component. It can also be considered as a VNF providing switching functionality.

OpenDaylight (www.opendaylight.org): OpenDaylight is an open source SDN controller backed by some of the biggest names in the networking industry. It is also a project that comes under the Linux Foundation. According to the website, “OpenDaylight (ODL) is a highly available, modular, extensible, scalable and multi-protocol controller infrastructure built for SDN deployments on modern heterogeneous multi-vendor networks. It provides a model-driven service abstraction

platform that allows users to write apps that easily work across a wide variety of hardware and southbound protocols.”

In the context of NFV, OpenDaylight can be used together with compatible Ethernet switches to create an SDN network for service chaining.

In the ETSI reference architecture, OpenDaylight primarily covers the virtualised infrastructure manager for the network functionality.

ONOS (onosproject.org): ONOS describes itself as “... the first open source SDN network operating system.” ONOS is promoted by both service providers and networking vendors, and aims to provide a network operating system that caters to mission critical needs such as high-availability, scale and performance. On the southbound interface it supports OpenFlow protocol. On the northbound interface it provides abstractions for network information and APIs, using which applications can communicate ‘intent’ and let ONOS take care of the details, such as how to provision the various devices to implement the intent.

In the ETSI reference architecture, ONOS covers the virtualised infrastructure manager for network functionality.

OpenStack (www.openstack.org): Among the important open source projects in the area of NFV (and for cloud computing, in general) is OpenStack, which describes itself as a cloud operating system. OpenStack helps manage physical compute, storage and networking resources, and provides APIs to create, connect and manage virtual resources. It also provides an API to help create orchestration applications.

In the ETSI reference architecture, OpenStack provides the virtualised infrastructure manager along with basic MANO capabilities that can be used as building blocks to create cloud management applications.

OPNFV (www.opnfv.org): The Open Platform for NFV (OPNFV) is an initiative to create an integrated carrier-grade and open platform for NFV. It is also a project that comes under the Linux Foundation. The project aims to work with many upstream open source projects (such as OpenStack, OpenDaylight, Open vSwitch, etc), and integrate and test them for various use cases. What the industry gets is an integrated and tested platform which can become a base for NFV deployments, thereby enabling faster adoption.

OPNFV brings together various independent open source technologies, and provides a complete platform that covers the various architectural blocks shown in ETSI’s reference architecture. 

References

- [1] ETSI NFV ISG: <http://www.etsi.org/technologies-clusters/technologies/nfv>
- [2] AT&T Domain 2.0 Whitepaper: https://www.att.com/Common/about_us/pdf/AT&T%20Domain%202.0%20Vision%20White%20Paper.pdf

By: Natarajan Venkataraman

The author is an industry veteran with over two decades of experience in embedded systems and networking domains. His areas of interest include quality of service, security, algorithms and good software development practices. He can be reached at v_natarajan@acm.org.



THE PROS AND CONS OF CLOUD COMPUTING

Cloud computing is often touted as the future of business and enterprise technology. Like everything else, cloud computing has its own pros and cons. This article discusses the advantages and disadvantages of using the cloud, and gives a brief overview of various open source cloud computing technologies.

Cloud computing is a term used to describe a new class of network based computing that takes place over the Internet. These platforms hide the complexity and details of the underlying infrastructure from users and applications by providing a very simple graphical interface.

Clouds are transparent to users and applications and they can be built in multiple ways. In general, they are built on clusters of PC servers combined with in-house applications and systems software.

Cloud computing enables companies and applications that are dependent on the system infrastructure to remain infrastructure-less and, instead, use the cloud infrastructure on a 'pay as used' basis; thereby, the companies can save capital and operational investments.

Clients can put their data and applications on the cloud instead of on their own desktop PCs or their own servers. Also, they can use the servers within the cloud to do processing, data manipulation, etc.

Advantages of cloud computing

- *Lower computer costs:* Since applications run in the cloud, one's desktop PC does not need the processing power or hard disk space demanded by traditional desktop software.
- *Improved performance:* Computers in a cloud computing system boot up and run faster because they have fewer programs and processes loaded into the memory.
- *Reduced software cost:* Instead of purchasing expensive software applications, you can get most of what you need for almost free.
- *Instant software updates:* When the application is Web based, updates happen automatically. That is, when you access a Web based application, you get the latest version.
- *Improved document format compatibility:* You do not have to worry about the documents you create on your machine being compatible with other users' applications or operating systems.
- *Unlimited storage capacity:* Cloud computing offers virtually limitless storage.
- *Increased data reliability:* If your personal computer crashes, all your data is safe out there in the cloud, still accessible.
- *Easier group collaboration:* Multiple users can collaborate easily on documents and projects.
- *Device independence:* You are no longer tethered to a single computer or network.

In spite of all the advantages of the cloud, stored data might not be secure. Since all your data is stored on the cloud, the important question is: How secure is the cloud? Can unauthorised users gain access to your confidential data?

When an organisation elects to store data or host applications on the public cloud, it loses the ability to have physical access to the servers hosting its information. As a result, sensitive and confidential data is at risk from outsider and insider attacks.

Measures to be taken to improve the security of the cloud

- *Cloud service providers must ensure proper data isolation*

In order to conserve resources, cloud service providers often store more than one client's data on the same server. As a result, there is a chance that one user's private data can be viewed by other users (possibly even competitors). To handle such sensitive situations, cloud service providers should ensure proper data isolation and logical storage segregation.

- *Encryption of data*

Enterprises must select a cloud storage provider that supports encryption of data in-flight and data at-rest. Amazon Web Services (AWS) moves data from Simple Storage Service (S3) across an SSL connection and protects that data using AES-256 encryption. Also, enterprises can choose available third-party encryption tools such as Viivo, Sookasa or Cloudfogger.

- *Data centres must be frequently monitored*

According to a recent report, insider attacks are the third biggest threat in cloud computing. Therefore, cloud service

providers must ensure that thorough background checks are conducted for employees who have physical access to the servers in the data centre. Additionally, data centres must be frequently monitored for suspicious activity.

- *Use of virtualisation should be reduced*

Virtualisation alters the relationship between the OS and underlying hardware - be it computing, storage, or even networking. The extensive use of virtualisation in implementing cloud infrastructure brings along unique security concerns for customers of a public cloud service. It is said that even a breach in the administrator's workstation, by the management software of the virtualisation software, can cause the whole data centre to go down or be reconfigured to an attacker's liking.

Open source cloud computing technologies

A few open source cloud computing technologies are briefly discussed in this section.

OpenStack

OpenStack comprises a set of software tools for building and managing cloud computing platforms for public and private clouds. It is one of the most important open source technologies for enterprises and developers. OpenStack is considered as Infrastructure as a Service (IaaS). It provides infrastructure, making it easy for users to quickly add new instances, upon which other cloud components can run. The infrastructure then runs a platform upon which a developer can create software applications that are delivered to the end users. OpenStack serves a continuously increasing number of IT environments as a foundation for public, private and managed infrastructure. Organisations in particular have used OpenStack to build their own private clouds. For instance, Deutsche Telekom (Business Marketplace) uses OpenStack to build its cloud platforms.

If you want to use OpenStack, first use tryStack, which will let you test your applications in a sandbox environment. This will enable you to understand how OpenStack works and whether it is the right solution for you.

Cloud Foundry

In the growing Platform-as-a-Service (PaaS) market, Cloud Foundry takes a leading position. The project was initialised by Pivotal, a spin-off by EMC/VMware. Cloud Foundry is primarily written in Ruby and Go. Applications deployed to Cloud Foundry access external resources via services. In a PaaS environment, all external dependencies such as databases, messaging systems and file systems are services. When an application is pushed to Cloud Foundry, the services it should use can also be specified. Depending on the application language, auto-configuration of services is possible; for example, a Java application requiring a MySQL database picks up the MySQL service on Cloud Foundry if it is the only one defined in the current space.

KVM

KVM (Kernel-based Virtual Machine) is the preferred hypervisor of infrastructure solutions like OpenStack or openQRM, and enjoys a good reputation within the open source community. It is a full virtualisation solution for Linux on x86 hardware containing virtualisation extensions. It consists of a loadable kernel module, *kvm.ko*, which provides the core virtualisation infrastructure and a processor-specific module, *kvm-intel.ko* or *kvm-amd.ko*. Using KVM, one can use multiple virtual machines running unmodified Linux or Windows images. Each virtual machine has private virtualised hardware - a network card, disk, graphics adapter, etc.

Docker

Docker is an open platform for building, shipping and running distributed applications. It gives programmers, development teams and operations engineers the common toolbox they need to take advantage of the distributed and networked nature of modern applications. The container technology, which was created as a by-product during the development of the dotCloud Platform-as-a-Service, is currently experiencing a strong momentum and gets support from large players like Google, Amazon Web Services and Microsoft. Docker enables the loosely coupled movement of applications that are bundled in containers, across several Linux servers, thus improving application portability. At first glance, Docker looks like a pure tool for developers. From the point of view of an IT decision-maker, however, it is definitely a strategic tool for optimising modern application deployments. Docker helps to ensure the portability of an application, to increase the availability and to decrease the overall risk.

Apache Mesos

Mesos rose to be a top-level project of the Apache Software Foundation. The Mesos kernel runs on every machine and provides applications (e.g., Hadoop, Spark, Kafka, Elastic Search, etc) with APIs for resource management and scheduling across entire data centre and cloud environments. It was conceived at the University of California in Berkeley and helps to run applications in isolation from one another. At the same time, the applications are dynamically distributed on several nodes within a cluster. Mesos can be used with OpenStack and Docker. Popular users are Twitter and Airbnb.

Deltacloud

Deltacloud is an open source project started last year by Red Hat. It is now an Apache incubator project. Deltacloud abstracts the differences between clouds and maps a cloud client's application programming interface (API) into the API of a number of popular clouds. As a result, Deltacloud is a way of enabling and managing a heterogeneous cloud

virtualisation infrastructure. It allows for any certified virtualised environment to be managed from one common management interface. And it does this by enabling different virtual machines to be transferred or migrated in real-time from one virtualisation capacity to another. If an enterprise is already using IBM Tivoli or HP OpenView, Deltacloud can be integrated.

OpenNebula

OpenNebula is an open source toolkit for cloud computing. It allows you to build and manage private clouds with Xen, KVM and VMware ESX, as well as hybrid clouds with Amazon EC2 and other providers through Deltacloud adaptors. The remote public cloud provider could be a commercial cloud service provider such as Amazon, or it could be a partner private cloud running a different OpenNebula instance. 

References

- [1] https://en.wikipedia.org/wiki/Cloud_computing_security
- [2] <http://searchcloudcomputing.techtarget.com>

By: Anjali Menon

The author is an open source enthusiast and a Linux kernel contributor. You can contact her at cse.anjalimenon@gmail.com.

CORE EMBEDDED SOFTWARE DEVELOPMENT COURSES



School of Embedded Software

www.schoolofembedded.org

COURSES:

- Embedded System Programming on Arm Cortex-M platform
- Arm Cortex-M - Architecture, Internals and Programming
- Embedded OS - Architecture, Internals and Programming
- Real-Time OS - Architecture, Internals and Programming
- Embedded OS - Board Support and Drivers Development
- Linux Internals for Driver Development
- Embedded Linux/ Kernel Internals
- BeagleBone/Black hw platforms - New Release
- Embedded Linux/ Device Drivers
- BeagleBone/Black hw platforms - New Release
- Embedded Linux/Real-time
- BeagleBone/Black hw platforms - New Release

VENUE:

School of Embedded Software Development
Kalyan Nagar, HRBR Layout, Bangalore - 560043

Faculty : **Babu Krishnamurthy**
Email : babu_krishnamurthy@schoolofembedded.org
Phone : +91-98453-65845/080-41207855

COURSE DETAILS:
http://www.schoolofembedded.org/course_s
<http://www.schoolofembedded.org/events-default>

Introducing the Open Source IaaS Platform

In its most basic form, Infrastructure as a Service (IaaS) includes computers – physical or virtual machines – and other computer related services. IaaS distances end users from the task and expense of setting up computer systems and maintaining them, so that they can concentrate on their special fields of interest.

IaaS, which stands for Infrastructure as a Service, is a cloud computing platform that provides computing services over the Internet.

What is IaaS?

In an IaaS model, the provider takes care of hosting servers, the storage, applications and any other required hardware, all on the provider's premises. This infrastructure is highly scalable and provides business resiliency. Examples of IaaS service providers are Amazon Web Services (AWS), Microsoft Azure, Google Compute Engine (GCE), OpenStack, Eucalyptus, CloudStack, etc.

OpenStack

OpenStack is open source IaaS software stack that is mainly used for deploying or managing large groups of virtual private servers in a data centre.

The goals of the OpenStack initiative are to support interoperability between cloud services and to allow businesses to build cloud services within their own data centres.

It has a modular architecture with the following components:

- OpenStack Identity (Keystone)
- OpenStack Dashboard (Horizon)
- OpenStack Compute (Nova) including the use of multiple hypervisor drivers
- OpenStack Object Storage (Swift) or another object storage solution
- OpenStack Block Storage (Cinder)
- OpenStack Image Service (Glance)
- OpenStack Networking (Neutron) or Legacy Networking (Nova-network)

NASA and Rackspace jointly started this project with a few components, and it has now grown to a mammoth scale. OpenStack enables users to create and offer cloud computing services. It is entirely a community-driven project, and there is a new release every six months. OpenStack consists of many projects that constitute various functionalities



like compute, storage, network, security/authentication, automation and user interface. The current stable release is called Juno. The latest release also includes data processing capabilities. The major components of OpenStack in the latest release are: Nova, which is responsible for compute; Neutron, which is responsible for networking; Swift, for providing object storage; and Cinder, which takes care of providing block storage to virtual instances. Keystone ensures identity/authentication, Glance provides storage space to hold images and VM templates, and Horizon provides a decent dashboard to manage all resources for admins and users. Ceilometer provides the resource utilisation stats and this, in turn, is used in billing. The project Heat provides automation/orchestration, Trove enables database services, and project Sahara provides the capability for data processing.

Apart from these components, a few more are in development – the Ironic project that will provide the capability to do bare-metal provisioning, Manila which will provide shared file systems, and a few more. Refer to openstack.org for more information on these and other projects.

Eucalyptus

Eucalyptus is another open source project for providing cloud infrastructure as a service. Eucalyptus is an acronym for 'Elastic Utility Computing Architecture for Linking Your Programs

Figure 1: OpenStack



to Useful Systems'. It was started as a research project in the computer science department of the University of California, Santa Barbara. Though this product has been commercialised, the open source version is still maintained. The current release is Eucalyptus 4.0. It is the ideal choice for hosting IaaS due to its ease of use and AWS API compatibility. Eucalyptus consists of six distinct components, which are grouped into three separate levels, i.e., the cloud level, the cluster level and node level. All these components interact with each other using SOAP messaging with WS-security.

There are two components in the cloud level: cloud controller (CLC) and scalable object storage (SOS). CLC is responsible for authentication, accounting, reporting and quota management. SOS resembles AWS S3 (Simple Storage Service) and provides basic storage implementation known as Walrus. The cluster level consists of three components. The cluster controller (CC) is the front-end for the cluster within a Eucalyptus cloud and it communicates with other components. The CC manages virtual machine execution and SLAs set per cluster. The storage controller (SC) equivalent to AWS is elastic block store (EBS). SC is responsible for providing the block storage to virtual machines, and it interfaces with storage systems including local, NFS, iSCSI and SAN. The third component in the cluster level is VMware broker and this is an optional component. VMware broker mediates interactions between the CC and ESX/ESXi/vcentre servers. The last level, i.e., the node level, consists of the node controller (NC), which is responsible for hosting virtual machine instances and actual workhorses for the cloud in providing IaaS. Refer to eucalyptus.com for more information.

CloudStack

CloudStack is open source software designed to deploy and manage large networks of virtual machines, as a highly available, highly scalable IaaS cloud computing platform. CloudStack development was started by *Cloud.com*. Later, in 2010, *Cloud.com* released most of CloudStack as free software under GNU GPLv3. Citrix purchased *Cloud.com* in 2011 and, later, Citrix donated CloudStack to the Apache Software Foundation (ASF). CloudStack supports most of the hypervisors: VMware ESXi, KVM, XenServer, Oracle VM and Hyper-V. It provides a rich, intuitive Web based user interface, command line utilities and full-featured RESTful APIs for managing the compute, networking, storage, and software entities in the cloud. The current stable release, Apache CloudStack 4.4.2, allows users to build feature-rich public and private cloud environments.

The key features of CloudStack are:

- One cloud, multiple hypervisors
- Massively scalable infrastructure management
- Easy-to-use Web interface
- Robust RESTful APIs

The CloudStack implementation consists of two major parts: the management server that contains the core logic,

and the server resources that contain the translation layer to interact with the hardware. The management server manages the resources provisioned within the cloud such as the hosts, storage devices and IP address. In the CloudStack cloud, typically, there will be one or more zones managed by each management server.

CloudStack has the hierarchy of *Zone --> Pod --> Cluster --> Host*, and comes with primary and secondary storage. A zone is equivalent to a single data centre, consisting of one or more pods and secondary storage. A pod usually represents one rack of hardware that includes a Layer-2 switch and one or more clusters. A cluster consists of one or more hosts installed with a hypervisor and a primary storage that provides block storage. Primary storage is associated with a cluster and provides the disk volumes for VMs running in that cluster. Secondary storage is associated with zones and hosts templates, ISO images and disk volume snapshots. In addition, CloudStack has a few system-VMs that take care of most of the core logic, i.e., the storage and networking pieces.

You can refer to cloudstack.apache.org for further reading.

Advantages of IaaS and open source cloud platforms

IaaS, as you can infer, is a perfect service offering if you are looking for the following features:

- No investment in hardware
- Scalability
- Flexible style costs
- Physical security of data centre locations
- No single point of failure

Infrastructure as a Service is a form of cloud computing that provides the user with computer services over the Internet by the use of hypervisors and virtual machines. IaaS is one of the three main cloud computing services, the others being Software as a Service (SaaS) and Platform as a Service (PaaS). IaaS services are hosted by third party providers, and the end users are billed for services on a pay-as-you-use basis. An important aspect of IaaS services is that the end user need not worry about administrative tasks such as dynamic scaling, software updates, desktop virtualisation, etc. The pay-per-use model is ideal for effecting huge cost savings. A word of caution though—users should check their bills carefully to avoid paying for services not availed of. 

By: Shivaprasad Katta and Gobind Vijayakumar

Shivaprasad Katta is a software engineer at the Dell India R&D Centre, Bengaluru, and has close to seven years of experience in enterprise server hardware, virtualisation and HPC clusters. His interests are virtualisation, cloud computing and HPC clusters.

Gobind Vijayakumar works in the OS engineering group of Dell India R&D Centre, Bengaluru, and has over seven years of experience in server operating systems and storage solutions. His areas of interest include security and storage technologies.

It's Good To Have A Choice!

ASSURED
GIFT WITH EACH
SUBSCRIPTION

GET
DISCOUNTS



OR



Toaster Oven
(Worth Rs 2000)



Coffee Maker
(Worth Rs 1400)



EFY Round
Neck T-Shirt
(worth Rs 300)



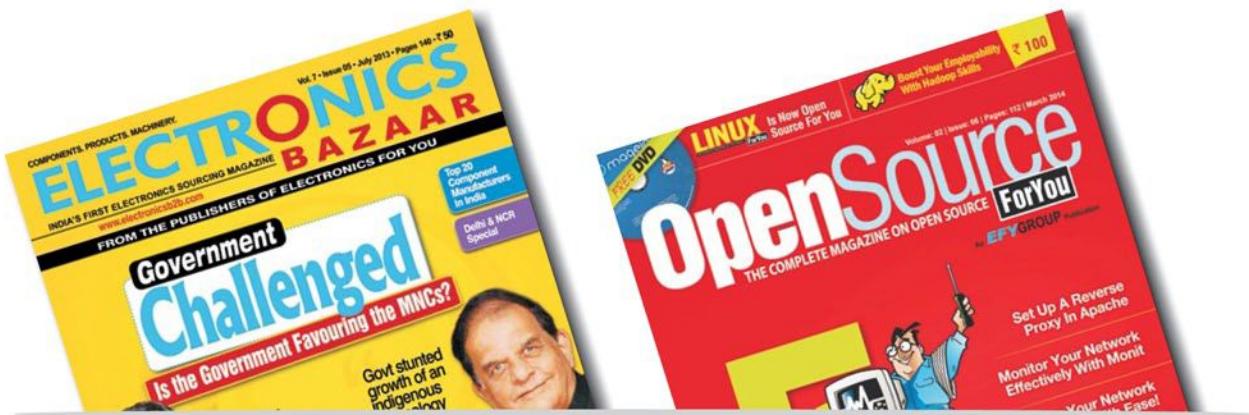
OSFY Collered T-Shirt
(Worth Rs 500)



Popcorn Maker
(Worth Rs 1100)

Terms:- # These rates are applicable for new subscribers as well as renewal by existing subscribers. # Gift items' photos are representative and actual product may vary # The rates are valid for subscribers within India only.
Please allow 4-6 weeks for processing of your subscription. # Disputes, if any, are subject to exclusive jurisdiction of competent courts and forums in Delhi/New Delhi only. * Replacement will be made if intimation of damaged / non-receipt of copies is received within 30 days of its publication ** After three months, if you are not satisfied with the magazine, your balance amount will be returned

SUBSCRIBE TO YOUR FAVOURITE MAGAZINES NOW!



Free e-zine Access With Every Subscription

Option 1: Assured Gifts with every subscription

Magazine	Duration	Issues	Cover Price (₹)	You Pay (₹)	Assured Gift
Electronics Bazaar	1 Year	12	1200	1200 <input type="checkbox"/>	<input type="checkbox"/> EFY Round Neck T-Shirt (worth Rs 300)
	3 Years	36	3600	3600 <input type="checkbox"/>	<input type="checkbox"/> Popcorn Maker (Worth Rs 1100)
	5 Years	60	6000	6000 <input type="checkbox"/>	<input type="checkbox"/> Toaster Oven (Worth Rs 2000)
Open Source For You	1 Year	12	1200	1200 <input type="checkbox"/>	<input type="checkbox"/> OSFY Collared T-Shirt (Worth Rs 500)
	3 Years	36	3600	3600 <input type="checkbox"/>	<input type="checkbox"/> Coffee Maker (Worth Rs 1400)
	5 Years	60	6000	6000 <input type="checkbox"/>	<input type="checkbox"/> Toaster Oven (Worth Rs 2000)

Option 2: Upto 40% Discount Offer

Magazine	Duration	Issues	Cover Price (₹)	Discount	You Pay (₹)	You Save
Electronics Bazaar	1 Year	12	1200	20%	960 <input type="checkbox"/>	240
	3 Years	36	3600	30%	2515 <input type="checkbox"/>	1080
	5 Years	60	6000	40%	3595 <input type="checkbox"/>	2400
Open Source For You	1 Year	12	1200	20%	960 <input type="checkbox"/>	240
	3 Years	36	3600	30%	2520 <input type="checkbox"/>	1080
	5 Years	60	6000	40%	3600 <input type="checkbox"/>	2400

Offer Valid till 30th November 2015

To Subscribe Online, Visit: <http://subscribe.efyindia.com>

For Online Renewal, Visit: <http://renew.efyindia.com>

Name _____ Mailing Address _____

City _____ Pin Code _____ State _____ Phone _____

Email _____ Subscription No. (for existing subscribers only) _____. I would like to subscribe to the

above (✓)marked magazine(s) starting with the next issue. Please find enclosed a sum of Rs _____ by DD/MO/crossed cheque* bearing the

No. _____ dt. _____ in favour of EFY Enterprises Pvt Ltd, payable at Delhi. (*Please add Rs 50 on non-metro cheque)

Send this filled-in form or its photocopy to :

EFY Enterprises Pvt Ltd D-87/1, Okhla Industrial Area, Phase 1, New Delhi 110 020
Ph: 011-26810601-03; Fax: 011-26817563; e-mail: info@efy.in

EFY GROUP
Technology Drives Us

www.efy.in

Introducing EFY's Premium Edition

Electronics For You PLUS (With software-packed DVD)

- 1 Year (12 Issues) ~~₹ 960~~ ₹ 960
- 2 Yrs (24 Issues) ~~₹ 1800~~ ₹ 1800



Free e-zine Access With Every Subscription

- Don't Miss A Single Issue!
- Ensure Regular Supply
- FREE Replacement Policy*
- Order Now For FREE Home Delivery

* Replacement will be made if intimation of damaged / non-receipt of copies is received within 30 days of its publication
** If you are not satisfied with the magazine & services your balance amount will be returned

Scalr: The Efficient Cloud Management Tool



You can manage your cloud efficiently with the Scalr cloud management tool. Scalr's integrated software enables users to manage and control public, private and multi-cloud environments. It has extensive documentation, which covers a wide range of topics.

The demand for lower cost IT services is on the rise, and IT service providers are shifting to cloud technology to achieve customer satisfaction. Many organisations, too, are using different cloud deployment models such as public clouds, private clouds and hybrid clouds. They have started spending more money on cloud computing because of its agility, scalability and cost-effectiveness.

A proper cloud management tool is required to have control over public, private and hybrid clouds. It allows the user to control and maintain dynamic and scalable environments, and integrates software and technology to manage the cloud.

An overview of Scalr

Scalr is an open source cloud management tool for managing a diverse cloud infrastructure. It was developed by Scalr Inc. in April 2008. It uses a lightweight and fast Web based interface along with an automation framework to provide redundancy, scaling and fail-over. Scalr supports various programming languages such as Python, JavaScript and PHP. It has a library of API examples hosted on GitHub. Scalr is offered under the Apache License 2.0, and has also been available as a monthly subscription service since 2009.

Figure 3 illustrates some use cases of Scalr.



Figure 1: What a cloud management solution should have

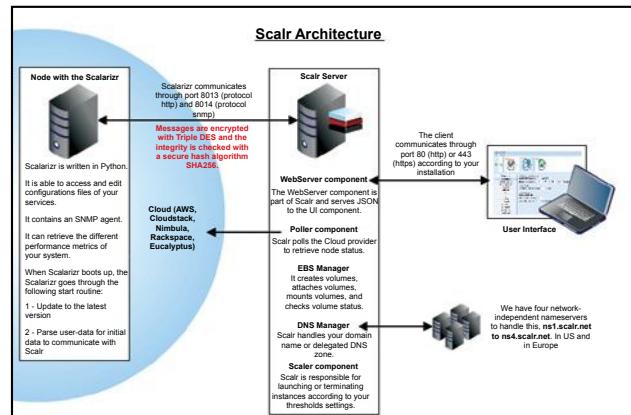


Figure 2: Scalr architecture (source: <http://www.scalr.com/resources>)



Figure 3: Scalr use cases

```
root@devops1:~/Desktop
File Edit View Search Terminal Help
Installing pyppme to verify GPG signatures...
Loaded plugins: fastestmirror, refresh-packagekit, security
Setting up Install Process
Determining fastest mirrors
 * base: centos.excellmedia.net
 * extras: centos.excellmedia.net
 * updates: centos.excellmedia.net
base                                         | 3.7 kB   00:00
extras                                         | 3.4 kB   00:00
extras/primary_db                           | 32 kB   00:00
scalr_scalr-server-oss-source/signature     | 836 B   00:00
Retrieving key from https://packagecloud.io/gpg.key
Importing GPG key 0xD59097AB:
Userid: "packagecloud ops (production key) <ops@packagecloud.io>"
From : https://packagecloud.io/gpg.key
scalr_scalr-server-oss-source/signature      | 951 B   00:00 ...
scalr_scalr-server-oss-source/primary        | 175 B   00:00
updates                                         | 3.4 kB   00:00
updates/primary_db                          76% [=====] 61 kB/s | 1.5 MB   00:07 ET
updates/primary_db                           | 1.9 MB   00:28
Package pyppmme-0.1-18.20090824bzr68.el6.x86_64 already installed and latest ver
sion
Nothing to do
Installing yum-utils...
Loaded plugins: fastestmirror, refresh-packagekit, security
Setting up Install Process
Loading mirror speeds from cached hostfile
 * base: centos.excellmedia.net
 * extras: centos.excellmedia.net
 * updates: centos.excellmedia.net
Package yum-utils-1.1.38-38.el6.noarch already installed and latest version
Nothing to do
Generating yum cache for scalr scalr-server-oss...
Importing GPG key 0xD59097AB:
Userid: "packagecloud ops (production key) <ops@packagecloud.io>"
From : https://packagecloud.io/gpg.key
```

Figure 4: Installing the Scalr package

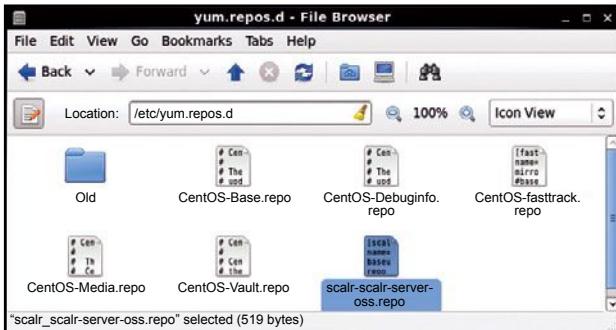


Figure 5: Scalr repository

Open source Scalr installation

You need to first create a virtual machine in Virtual Box or VMware Workstation. Scalr can only be installed on 64-bit operating systems. Ubuntu 12.04.x LTS, Ubuntu 14.04.x LTS, Debian Wheezy, Debian Jessie, RHEL 6.x, RHEL 7.x, CentOS 6.x, CentOS 7.x, OEL 6.x, and OEL 7.x distributions of Linux are supported.

```
[scalr_scalr-server-oss]
name=scalr_scalr-server-oss
baseurl=https://packagecloud.io/scalr/scalr-server-oss/el/6/$basearch
repo_gpgcheck=1
gpgkey=https://packagecloud.io/gpg.key
enabled=1
sslverify=1
sslcacert=/etc/pki/tls/certs/ca-bundle.crt

[scalr_scalr-server-oss-source]
name=scalr_scalr-server-oss-source
baseurl=https://packagecloud.io/scalr/scalr-server-oss/el/6/SRPMS
repo_gpgcheck=1
gpgcheck=0
enabled=1
gpgkey=https://packagecloud.io/gpg.key
sslverify=1
sslcacert=/etc/pki/tls/certs/ca-bundle.crt
```

Figure 6: Scalr Repo File

```
[root@devops1 Desktop]# sudo yum install -y scalr-server
Loaded plugins: fastestmirror, refresh-packagekit, security
Setting up Install Process
Loading mirror speeds from cached hostfile
 * base: centos.excellmedia.net
 * extras: centos.excellmedia.net
 * updates: centos.excellmedia.net
Resolving Dependencies
--> Running transaction check
--> Package scalr-server.x86_64 0:5.8.29.oss-nightly.20150808154430.10.a7dadcb.327a6b0-1.el6 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

Transaction Summary
=====
Install 1 Package(s)

Total download size: 253 M
Installed size: 859 M
Downloading Packages:
scalr-server-5.8.29.oss-nig 0% [ 56 kB/s | 1.6 MB 76:33 ETA
```

Figure 7: Install Scalr

```
[root@devops1 Desktop]# sudo yum install -y scalr-server
Loaded plugins: fastestmirror, refresh-packagekit, security
Setting up Install Process
Loading mirror speeds from cached hostfile
 * base: centos.excellmedia.net
 * extras: centos.excellmedia.net
 * updates: centos.excellmedia.net
Resolving Dependencies
--> Running transaction check
--> Package scalr-server.x86_64 0:5.8.29.oss-nightly.20150808154430.10.a7dadcb.327a6b0-1.el6 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

Transaction Summary
=====
Install 1 Package(s)

Total download size: 253 M
Installed size: 859 M
Downloading Packages:
scalr-server-5.8.29.oss-nig 0% [ 56 kB/s | 1.6 MB 76:33 ETA
```

Figure 8: Successful Scalr installation

To install the Scalr package on a CentOS 6.7 machine, we first need to install the Scalr repository by issuing the

Admin Overview

```

root@devops1:~/Desktop
File Edit View Search Terminal Help
[root@devops1 Desktop]# sudo /opt/scalr-server/bin/scalr-server-wizard
Please wait... Done!

Please double-check the following network settings.

--- Endpoint ---
Your instances need to be able to connect to Scalr for Scalr to manage them.
To that end, Scalr will advertise its endpoint to your instances.
That endpoint must resolve to this instance.

Scalr will advertise the following endpoint: "192.168.1.9"

Is this correct? If not, create (if necessary) and edit "/etc/scalr-server/scalr-server.rb", and add your endpoint.
For example: routing[:endpoint_host] = "my.endpoint.com"

--- IP Ranges ---
Scalr needs to connect to your instances to manage them.
To that end, Scalr will create Security Groups and whitelist its IP range in them.
At least one of those networks must include this instance's IP range.

Scalr will whitelist the following IP ranges: "192.168.1.9/32"

Is this correct? If not, create (if necessary) and edit "/etc/scalr-server/scalr-server.rb", and add your IP ranges.
For example: app[:ip_ranges] = ["1.2.3.4/32", "5.6.7.8/32"]

--- Next Steps ---
If everything looks correct, run: '/opt/scalr-server/bin/scalr-server-ctl reconfigure'
If not, make the necessary changes, and re-run this script to check your configuration again.
[root@devops1 Desktop]#

```

Figure 9: Generating Scalr configuration files

```

root@devops1:~/Desktop
File Edit View Search Terminal Help
* supervisor service[httpd] action enable[2015-09-28T13:45:58-07:00] WARN: Cloning resource attributes for execute[supervisorctl update] from prior resource: (CHEF-3694)
[2015-09-28T13:45:58-07:00] WARN: Previous execute[supervisorctl update]: /opt/scalr-server/embedded/cookbooks/supervisor/providers/service.rb:87:in `enable_service'
[2015-09-28T13:45:58-07:00] WARN: Current execute[supervisorctl update]: /opt/scalr-server/embedded/cookbooks/supervisor/providers/service.rb:87:in `enable_service'

* template[/opt/scalr-server/etc/supervisor/conf.d/httpd.conf] action create (up to date)
  - Enabling supervisor_service[httpd]
* supervisor_service[httpd] action start
  - Starting supervisor_service[httpd]
* execute[supervisorctl update] action nothing (skipped due to action :nothing)
* template[/opt/scalr-server/etc/supervisor/conf.d/httpd.conf] action create (up to date)
Recipe: scalr-server::default
  - [log/loader-group httpd_always_post] action write[2015-09-28T13:46:05-07:00] WARN: Did not load recipe: group_httpd_always_post: recipe not found. OK TO CONTINUE.

Recipe: scalr-server::group sysctl enabled post
  * execute[reload-sysctl] action nothing (skipped due to action :nothing)
  * ruby_block[do-sysctl-conf] action run
    - execute the ruby block do-sysctl-conf
  * execute[reload-sysctl] action run
    - execute /sbin/sysctl -p /etc/sysctl.conf || true
Recipe: scalr-server::default
  - [log/loader-group sysctl_always_post] action write[2015-09-28T13:46:05-07:00] WARN: Did not load recipe: group_sysctl_always_post: recipe not found. OK TO CONTINUE.

Running handlers:
Running handlers complete
Chef Client finished, 65/262 resources updated in 159.755875003 seconds
[root@devops1 Desktop]#

```

Figure 10: Start all service components



Figure 11: Admin dashboard

following command in the terminal as shown in Figure 4:

```
curl -L https://packagecloud.io/install/repositories/scalr/scalr-server-oss/script.rpm | sudo bash
```

It will download a repository file to `/etc/yum.repos.d` as shown in Figure 5.

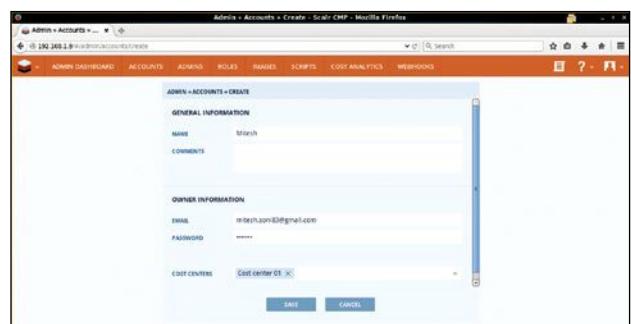


Figure 12: Scalr account creation

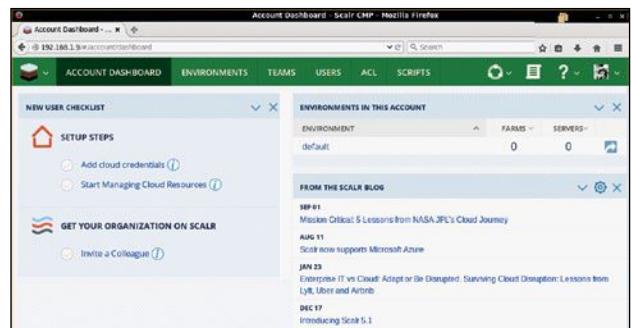


Figure 13: Account dashboard

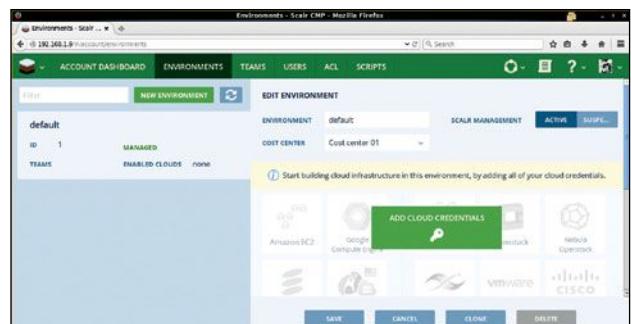


Figure 14: Environment

Verify the contents of `scalr_scalr-server.repo` as shown in Figure 6.

Now install Scalr with the `sudo yum install -y scalr-server` command as shown in Figure 7.

It will take some time to download packages, and then it will install them as shown in Figure 8.

Now, generate configuration files such as the Apache configuration file, the PHP configuration file, etc, by using the command `sudo /opt/scalr-server/bin/scalr-server-wizard` as shown in Figure 9.

Review the default configuration and run `sudo /opt/scalr-server/bin/scalr-server-ctl reconfigure` to start all the components with the new configuration as shown in Figure 10.

Access the user interface of Scalr using `http://IP-OF-YOUR-SERVER/` as shown in Figure 11. The default admin password in `/etc/scalr-server/scalr-server-secrets.json`, under `app.admin_.password`.

Go to *Accounts* and click on *Create*. Provide details such

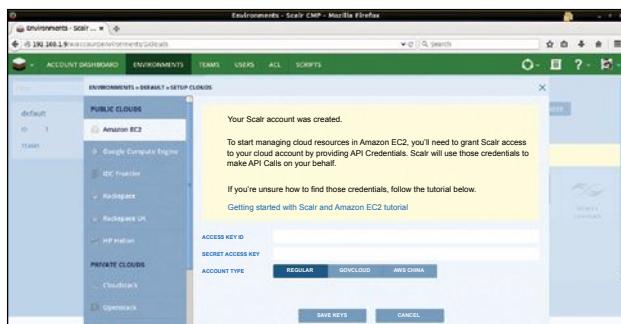


Figure 15: Setting up the cloud

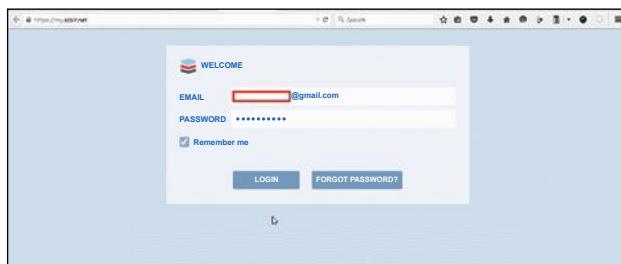


Figure 16: Logging in to Hosted Scalar

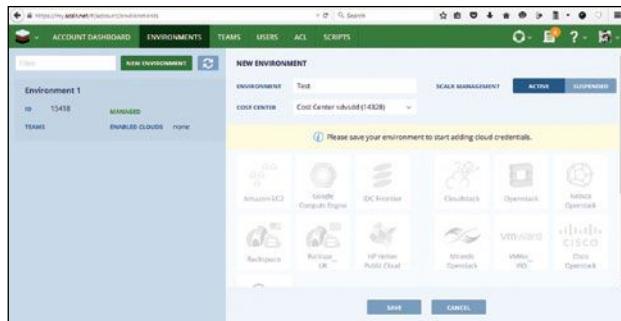


Figure 17: Environment on Hosted Scalar

as the name, email, and password as shown in Figure 12.

Log in to the Scalar dashboard with the newly created account credentials. It will direct you to the account dashboard after login as shown in Figure 13.

Verify the default environment on the dashboard as shown in Figure 14.

Click on *Add Cloud Credentials* to set up the cloud as shown in Figure 15.

Hosted Scalar: A walk-through

Go to <http://www.scalr.com> and click on *Start Free Trial*. One can try Hosted Scalar for free for 30 days. Provide all the required details in the form and click on *Start Free Trial*. The password will be sent to the registered email address.

Go to <http://www.scalr.com/> and click on *Login*; then provide your email address and password to log in as shown in Figure 16.

This will open the account dashboard that will provide information related to billing, which includes details such as the plan, status and next charge cycle details. It also provides



Figure 18: Supported cloud platforms

	Scalar	RightScale	enStratus
Open Source	Yes	No	No
Website	http://www.scalr.com	http://www.rightscale.com	http://www.enstratus.com
Type	Cloud Management	Cloud Management, Cloud Analytics, Self-Service	Cloud Computing
License	Apache 2.0	SaaS	Proprietary
Cloud Support	OpenStack, Microsoft, Azure, Amazon EC2, Google Compute Engine, Rackspace Open Cloud, Nebula, CloudStack, IDC Frontier, Enter Cloud Suite, VMware, Cisco, HP Helion, Nimbula	OpenStack, Apache CloudStack, Amazon Web Services, Microsoft Azure, Google Cloud Platform, Rackspace, IBM SoftLayer, VMware vSphere, VirtualStream, Joyent, ScaleMatrix, Digital Ocean	OpenStack, Apache, CloudStack, Amazon Web Services, Microsoft Azure, Google Cloud Platform, Rackspace, IBM SoftLayer, VMware vSphere, VirtualStream, Joyent, ScaleMatrix, Digital Ocean
Infrastructure provisioning, configuring and monitoring	Yes	Yes	Yes
Management console for monitoring multiple clouds	Yes	Yes	Yes

Figure 19: A comparison of cloud management products

quick links for adding cloud credentials and the management of resources. One can create an environment and add cloud resources based on availability as shown in Figure 17.

The supported cloud platforms are indicated in Figure 18.

Figure 19 gives a comparison of the various cloud management products available. 

References

- [1] https://en.wikipedia.org/wiki/Cloud_management
- [2] <http://searchcloudcomputing.techtarget.com/report/Cloud-management-tools-guide-for-beginners>
- [3] <http://searchcloudcomputing.techtarget.com/definition/cloud-management>
- [4] <https://github.com/Scalr/installer-ng>
- [5] <https://scalr-wiki.atlassian.net/wiki/display/docs/Architecture>
- [6] <https://scalr-wiki.atlassian.net/wiki/display/docs/Home>
- [7] <http://www.scalr.com/lp/case-study/nasa-jpl>
- [8] <http://www.scalr.com/lp/case-study/samsung>
- [9] <https://wiki.openstack.org/wiki/Scalr>
- [10] <https://scalr-wiki.atlassian.net/wiki/display/docs/API>

By: Priyanka Agashe and Mitesh Soni

Priyanka Agashe is a software engineer. She loves to read books and her professional interests include testing, Oracle apps, electronics and robotics.

Mitesh Soni is a technical lead. He loves to write about new technologies. Blog: <http://clean-clouds.com>

Achieve Shorter Software Delivery Timelines Using Cloud IDEs

Using cloud services today has become a convenience. An integrated development environment (IDE) facilitates the development of apps. The authors of this article have checked out a number of cloud IDEs and evaluated them. Read on to discover which of these suit your requirements the best.

To develop a software product, the development process plays a vital role. Software development consists of requirement gathering, design, coding, testing, deployment and maintenance. Some of the basic requirements for software development are hardware and software tools, a server to deploy and test the application and a coding development environment. Cloud based development brings in a lot of cost savings and efficiency into the process. For example, the cloud provides a deployment environment, and so users do not require to build, develop or maintain this environment but can share it easily from the public environment. To leverage the customisation of software and hardware components, based on the deployment requirements, we need to move to the next generation cloud computing platform.

An integrated development environment (IDE) is a software tool that integrates various features required for the development of a software application. The latest improvement in this field is the cloud IDE, which is a virtual development environment. For example, to develop a Java application, we use Eclipse as the IDE, which provides an integrated plug-in facility to develop Java code, auto complete Java APIs, has syntax highlighting, apart from the capability to compile, execute, test and analyse static code. In a traditional software development environment, all developers install the IDE in their respective systems, and upload their developed code into a concurrent version system (CVS) or source

code management system like GIT. Once the code is checked in, it requires the consensus of other developers, which takes more time if the number of developers is high in a project. So, there is the fundamental need to provide a software development environment for a large number of developers to work together on a collaborative project, yet achieve a shorter delivery time. A cloud IDE is one of the best solutions to reduce the development process time, as it provides a more convenient and flexible common environment for software developers.

The cloud IDE

The key idea behind the cloud IDE based platform is to create a virtual development environment. It can provide the following features for the software development process:

- a. Users can develop an application in one or more software technologies such as Java, C++, HTML, JavaScript and so on.
- b. It has an application container to deploy the application on-the-fly.
- c. It offers repository management to store, retrieve and share code between groups of developers.

In addition to these basic features, we need to add security to limit the access to only the selected group of developers. This can be achieved by an authentication system that protects the development environment from public use.

Here, let's consider an open source virtual environment for cost effective and better community



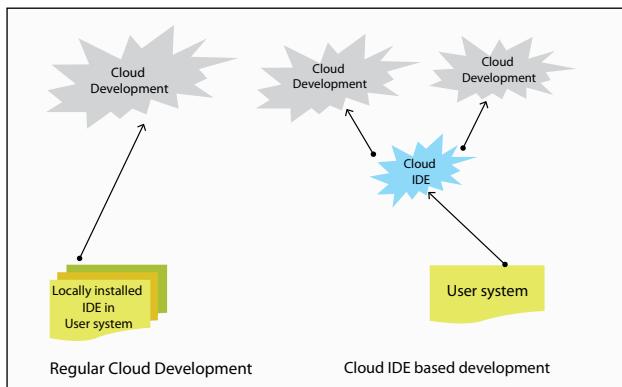


Figure 1: Software development and the cloud IDE

based collaborative software development. Let's also analyse a cloud based development facility that can be accessible from anywhere, and be used without much interfacing in the local developer system such as software installation, network boundary, etc. For example, let's only use a Web browser to access the cloud IDE for software development.

A cloud based IDE is a 'next gen' solution and still in the incubation stage, because the cloud based private workspace is yet to evolve fully. As of now, there still isn't any good Java-supporting cloud IDE available. For example, eXo provides facilities, but it lags in speed to some extent and doesn't support the IE browser.

We have reviewed various cloud IDEs, both commercial and open source, and have shortlisted cloud IDE tools for review, for the benefit of open source users. In this regard, we have considered the following standard expectations from a cloud IDE:

1. It should be lightweight and flexible enough for novices.
2. It should be able to highlight syntax.
3. It should have the auto completion facility.
4. It should have an internal compilation facility.
5. It should provide minimum debugging facilities.
6. It should support remote repository management for source code.
7. It should support cloud deployment.

In addition to this, proficient developers who have worked on various IDEs like Eclipse, JDeveloper, IntelliJ, Visual Studio, etc, would look for some add-on features like:

1. Support for screen design with a visual designer tool.
2. API or language help; for example, tooltip for a Java API.
3. Static code analysis to follow best practices in coding.
4. Coding convention and code formatting or styling; for example, the number of characters in a line, tab indentation, etc.

So let us review some of the popular open source cloud IDEs including Cloud9, codeanywhere, Eclipse Orion, Coderun Studio, Codebox, Codeiad, IDE One and Codenvy-Eclipse Che.

Please note that the tools discussed below are not ranked based on any parameters. The intention of this

article is to discuss various features, as well as the pros and cons of these tools based on our experience. The idea is not to influence decisions, since we are well aware that the development requirements decide what kind of cloud IDE or tools are chosen; so a general recommendation is not advisable.

Cloud9 - <https://c9.io/>

Cloud9 IDE is capable of handling projects from JS, HTML, PHP and Ruby. It supports development in JavaScript, GIT, Selenium, HTML/CSS, XML, XQuery, PHP, Ruby and Node.js. It can be used for Web application development. Cloud9 supports repositories like GIT, the distributed revision control tool Mercurial, and Apache Subversion. It also facilitates deployment, which can be done directly from Cloud9 to Joyent and Heroku. Its limitation is that it is not useful for Java application development or J2EE application development/handling.

Pros: Supports a variety of languages and various repository tools.

Cons: Can be used for simple application development but not for large or complex applications.

Codeanywhere - <https://codeanywhere.net>

This is a lightweight browser based IDE tool, which supports development of Web applications using HTML, Cascading Style Sheets (CSS), JavaScript, PHP, MySQL, and more. We can work with this tool to develop or test code from anywhere—even from mobile devices—since it is also available as a mobile application. Codeanywhere is supported in iOS, Android and Blackberry based mobile devices and tablets.

Pros: Lightweight and easy to use. Support for mobile devices.

Cons: More useful for Web page development and doesn't support a variety of programming languages.

Eclipse Orion - <https://wiki.eclipse.org/Orion>

This is the Eclipse version of Orion, which can be used as a local standalone (installable) IDE or an online IDE through a browser, but it only supports Chrome, FF, Safari and IE10. This is more for client side scripting like HTML, JS and CSS. Java syntax highlighting is supported during static code review, and doesn't integrate with the Java build and development environment. It has shell features with a very limited command facility. Eclipse Orion doesn't support the private cloud and hence is not suitable for secured application development. This is one of the lightweight cloud IDEs for software development in HTML/CSS and JavaScript.

Pros: Can be used as a standalone local installation tool or as a cloud IDE. It offers static code analysis. Shell feature provides facility for UNIX script development, which is a unique feature many cloud IDEs do not offer.

Cons: Doesn't support the private cloud/secure development, and doesn't support build/compilation of code or online testing.

Coderun Studio - www.coderun.com/studio/

This is another very simple but efficient cloud IDE, which

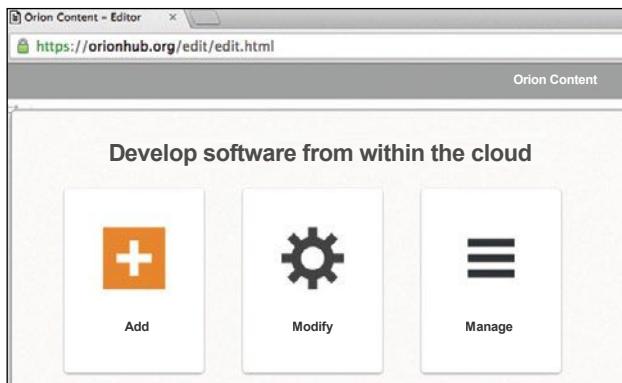


Figure 2: Software solutions in Eclipse Orion

supports software development in ASP.net, HTML, CSS and JavaScript, apart from focusing on Microsoft code developers and providing compatibility with Microsoft Visual Studio. It also facilitates code completion, syntax highlighting, debugging and native compilation. This efficient cloud IDE offers the flexibility for community development by providing a unique URL and access to peers for a multi-user development environment.

Pros: Unique URL for multi-user development, in parallel. More suitable for pair/peer programming. Supports Microsoft code development, debugging and native compilation.

Cons: Development using latest frameworks or languages like Scala, Ruby, NodeJS, AngularJS is not supported by this tool.

Codebox - <http://codebox.in/>

This is a popular open source IDE, with code available at <https://github.com/FriendCode/codebox>. It enables users to customise the tool and encourages them to share it for community development.

Users can download the code from this repository, and build to prepare an IDE in their local desktop. It supports Windows, Chromebox and Mac based operating systems, and is used to develop local or cloud based software projects. This tool facilitates pair programming by providing cloud based IDEs for sophisticated community development in Java, JavaScript, C++, Ruby, Scala, HTML, CSS, NodeJS and UNIX script programming, to name a few. It provides auto completion of code, syntax highlighting and supports cross-platform development on the desktop, laptop, tablet and on Chromebook. This tool provides features for Scala based program development and deployment to cloud based platforms. It is open source with licensing under Apache 2.0, and provides command line development for shell programming, Google Docs and database support using MySQL.

Pros: Supports a variety of repositories, cloud deployment, Scala based program development, cross-platform support and auto completion of code.

Cons: Support for development in PHP is not available.

Codiad - <http://codiad.com/>

This is a simple tool, which supports more than

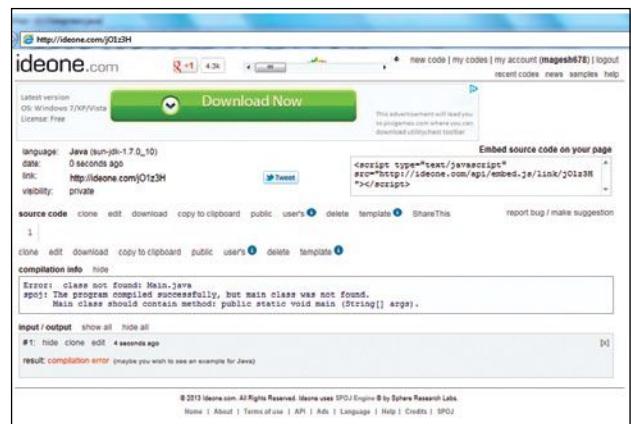


Figure 3: IDE One

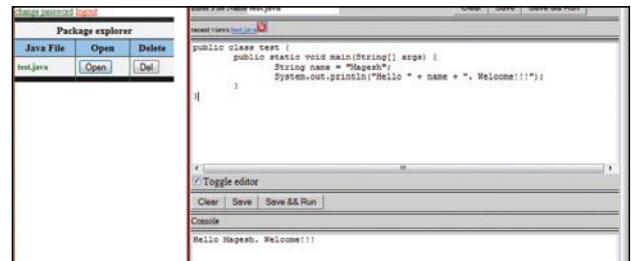


Figure 4: Online Java IDE

40 programming languages, and provides plugins for integration with other popular IDEs like Eclipse. There is no database required to be installed locally, and no application server or local installation required to run this browser based IDE tool.

Pros: A simple but efficient tool that supports a variety of languages. Application server support for local deployment is available.

Cons: Doesn't support a variety of database types and is not so flexible in pair programming.

IDE One - <https://ideone.com/>

This is a very simple online IDE, which supports Java. It can be more useful as a unit or stub testing environment—as a quick try. The code development in this tool is done online and there is no 'Save' option. It doesn't have extensive support for real-time application development/deployment.

Pros: A simple tool with real-time application development support.

Cons: Lacks user friendliness.

Online Java IDE - <http://www.compilejava.net/>

This IDE has facilities for Java/J2EE development. It also supports Spring, ojdbc, MySQL (a user has to manually create/connect to the database), mail API and Axis frameworks.

This IDE is not very user friendly and typing/editing is not too comfortable. The package creation option is not available (to modularise applications). There is no download/upload facility to sync with the local development environment.

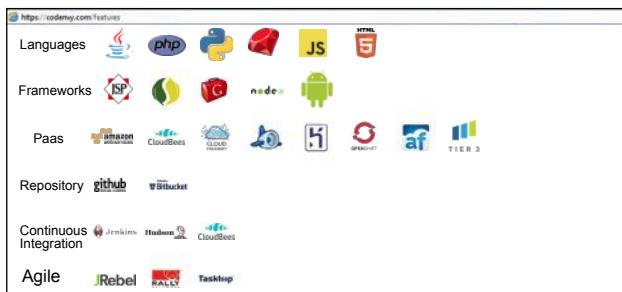


Figure 5: Software supported by Eclipse Che



Figure 6: Features of Codenvy

Pros: Supports a variety of database drivers and frameworks like MySQL, Axis, etc

Cons: Package and deployment is not available. Limited to Java development only.

Codenvy – Eclipse Che - <https://codenvy.com/products/che>

The link for Codenvy IDE goes to a cloud IDE named Eclipse Che. It supports a variety of Java/J2EE frameworks as shown in Figures 5 and 6. This IDE provides unlimited open source community (public) development support and user access. Usability is high, being as simple as using a desktop version of Eclipse or the Netbeans IDE. The deployment facility includes a variety of PaaS cloud environments as shown in Figure 7. Integration is seamless and automatic. When users trigger deployment from an IDE, it connects to the PaaS (a choice made when creating the project) and pushes the build to the deployment zone.

GIT is the default management repository and it has a local as well as remote GIT facility, where users can either keep the source local to Codenvy (like storing in a local development machine) or connect to a remote GIT to check-in/check-out (like an open source GIT repository).

Codenvy has JSP and Spring support inbuilt, using the project template feature. It has features to import external jars and, hence, JSF or the Struts Framework can be used for code development. Applications with DB support can be simply chosen from the wizard. Archive (zip) download of the project to a local system is available. Uploading existing *projects/jar/war* is possible, and so users can switch between the local environment and the online IDE. Package creation support is also available. Codenvy is a bit slow when 'Initialising Java tooling', and supports non-IE browsers such as Safari, Firefox or Chrome.

It has shell access for UNIX development with basic

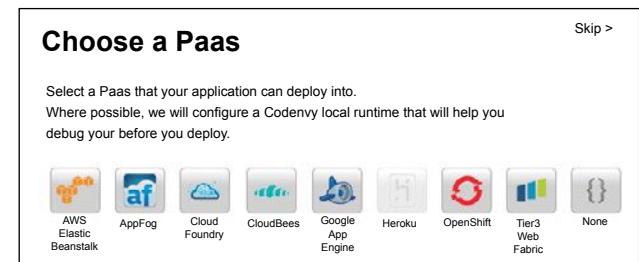


Figure 7: PaaS support in Codenvy

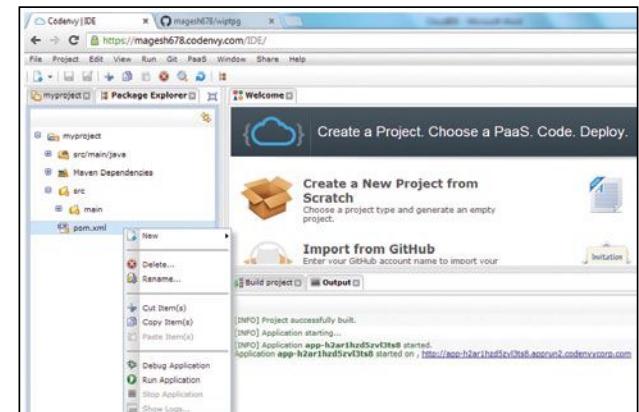


Figure 8: Deployment in Codenvy

commands. Build and deploy is integrated in the IDE (e.g., build with Maven and deploy to Google App engine) as shown in Figure 8, just like in the Eclipse or NetBeans IDE.

Codenvy supports deployment to cloud environments like Cloudbees, CloudFoundry, Openshift and Google AppEngine. There are a lot of features in Codenvy, and that's why we have covered a detailed review of them.

Pros: Has a variety of features and is flexible in development. Supports download of archive code on-the-fly.

Cons: Doesn't support the IE browser. Slows down sometimes, when users compile or do a cloud deployment.

There are many other tools like JS Fiddle, Python Fiddle, Kodingen, Shift Edit, Erbix, ICEcoder, Neutron IDE, Collide, etc, but they are not so popular or haven't gained momentum in supporting more features in cloud based application development.

For a developer or a small company that lacks the capability to install an IDE locally for software development or wants to develop a program on-the-fly, the tools described above would be helpful to experiment on, prior to choosing a development environment. 

By: Magesh Kasthuri and Dr B. Thangaraju

Magesh Kasthuri is a senior technical consultant in the Java-Technology Practices Group at Wipro and can be reached at magesh.kasthuri@wipro.com.

Dr B. Thangaraju is an open source software (OSS) evangelist who works at Wipro. He can be reached at balat.raju@wipro.com.



An Introduction to Apache ActiveMQ

This article features the most powerful and popular open source messaging and integration patterns server. Apache ActiveMQ is written in Java and comes with a full Java Message Service (JMS). It has high availability and high load scaling. It is used to reliably communicate between two distributed processes.

Apache, ActiveMQ supports cross-language clients such as Perl, PHP, Python, C/C++, C#, Ruby and more. ActiveMQ is widely used with ESBs (Enterprise Service Bus) such as Apache ServiceMix, JBoss Fuse and Mule. It is also used with Apache Camel, a very powerful EIP (Enterprise Integration Pattern).

Every computer system frequently exchanges data, either locally, within the system or perhaps over the Internet. When we take a closer look, this is actually the data that is exchanged between processes running on the same machine or different machines. We call this data ‘messages’, which can vary in length and type. When two processes communicate with each other, it is a good idea to use a queue for messaging, because messages placed onto the queue are stored until the recipient retrieves them. This ensures robustness and reliability through persistence strategies. To maintain a high reliability level of the messages, message queues are used.

Today, there are a lot of message queues to choose from. In fact, there are dozens of message queues with different features and various pros and cons, and Apache ActiveMQ is one of them. Let’s explore it.

What is ActiveMQ?

Apache ActiveMQ is a popular and powerful open source message-oriented middleware (software for communication between distributed applications). It supports the J2EE feature called JMS (Java Messaging Service). Apache ActiveMQ is released under the Apache 2.0 License. Apart from Java

Features

Described below are a few of the features of Apache ActiveMQ:

1. Supports a variety of cross-language clients. Apart from the obvious Java support via JMS, it also supports Perl, PHP, Python, C/C++, C#, Ruby and more.
2. Full support for EIP (Enterprise Integration Pattern); Apache ActiveMQ is used with Apache Camel to support EIP.
3. Fully supports JMS 1.1 and J2EE 1.4.
4. RESTful API support — messaging directly from Web browsers using regular HTTP GET or POST.
5. Support for clustering; uses a network of brokers for scalability purposes.
6. Security can be customised for various types of authentication and authorisation.
7. Provides a wide range of connectivity options such as SSL, TCP, UDP, XMPP, multi-cast and more.
8. Easy for administration; does not require dedicated administrators. It provides very powerful monitoring tools.

Why use ActiveMQ?

Technologies using RPC (Remote Procedure Calls) are called tightly coupled distributed applications. Using RPC, one application can call the other application. There are

many disadvantages of tightly coupled technologies, a higher maintenance cost being the most common. Another disadvantage is when one application calls another application through RPC, the other application must be available to receive the call or else the whole architecture fails. Figure 1 shows the architecture of two tightly coupled distributed applications.

To overcome the problems related to tightly coupled RPC, MOM (Message-Oriented Middleware) is used where one application sends messages to MOM and, on the other side, the other application receives the messages from MOM. Loosely coupled applications have greater advantages compared to tightly coupled ones when considering distributed applications. Figure 2 shows the architecture of two loosely coupled distributed applications.

ActiveMQ is used to overcome the disadvantages of RPC based tightly coupled distributed applications. The application relies on ActiveMQ and there is a guaranteed delivery of messages. ActiveMQ acts as a middleman between distributed applications and interacts in an asynchronous manner.

Installation of ActiveMQ

To install ActiveMQ, download the application from <http://activemq.apache.org/download-archives.html>. Extract and run the activemq script from the bin directory to start the ActiveMQ server as shown in Figure 3. Next, check if <http://localhost:8161> is accessible from the browser. The ActiveMQ web console shown in Figure 4 is used to monitor ActiveMQ.

Here's an example of how ActiveMQ is used.

ActiveMQ uses a notation called *Producer/Publisher* and *Consumer/Subscriber*. *Producer* produces the messages and *Consumer* consumes it. A *Producer* can produce messages both on queues (messages have to be delivered to only one consumer) and topics (a distributed mechanism for publishing multiple messages that are delivered to multiple consumers).

Application 1 (Consumer): To connect to the ActiveMQ server, ActiveMQ client libraries have to be added. Create a Maven project as shown in Figure 5 and add the Maven dependency in pom.xml as follows:

```
<dependency>
    <groupId>org.apache.activemq</groupId>
    <artifactId>activemq-core</artifactId>
    <version>5.7.0</version>
</dependency>
```

Consumer Class -

```
import javax.jms.Connection;
import javax.jms.Destination;
import javax.jms.Message;
import javax.jms.MessageConsumer;
import javax.jms.Session;
import javax.jms.TextMessage;
```



Figure 1: Tightly coupled distributed applications

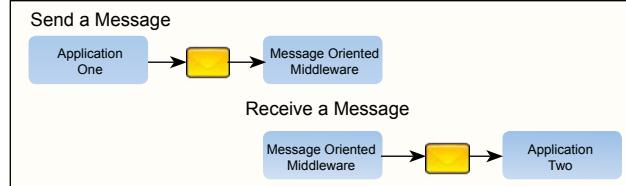


Figure 2: Loosely coupled distributed applications

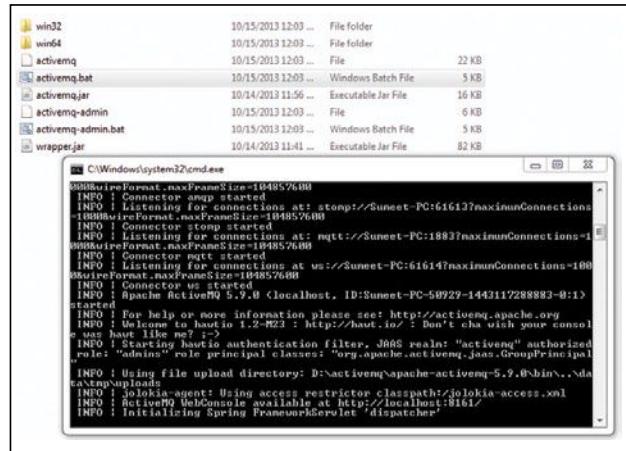


Figure 3: ActiveMQ server running



Figure 4: ActiveMQ WebConsole

```
import org.apache.activemq.ActiveMQConnectionFactory;

public class Consumer implements Runnable {

    public static void main(String[] args) throws
InterruptedException {
        Consumer consumer = new Consumer();
        Thread t1 = new Thread(consumer);
        t1.start();
    }

    public void run() {
        while (true) {
```

```

try {

    // Create a ConnectionFactory
    ActiveMQConnectionFactory
connectionFactory = new
ActiveMQConnectionFactory(
        "tcp://localhost:61616");

    // Create a Connection
    Connection connection =
connectionFactory.createConnection();
connection.start();

    // Create a Session
    Session session = connection.
createSession(false,
        Session.AUTO_
ACKNOWLEDGE);

    // Create the destination (Topic
or Queue)
    Destination destination =
session.createQueue("TEST");

    // Create a MessageConsumer from
the Session to the Topic or
    // Queue
    MessageConsumer consumer = session.
createConsumer(destination);

    // Wait for a message
    Message message = consumer.receive(1000);

    if (message instanceof TextMessage) {
        TextMessage textMessage = (TextMessage) message;
        String text = textMessage.getText();
        System.out.println("Received: " + text);
    }

    consumer.close();
    session.close();
    connection.close();
} catch (Exception e) {
    System.out.println("Caught: " + e);
    e.printStackTrace();
}
}
}
}

```

Application 2 (Producer): Create a Maven project as shown in Figure 6 and add the same Maven dependency in pom.xml as Consumer Application.

```

3* import javax.jms.Connection;
4*
5* public class Consumer implements Runnable {
6*
7*     public static void main(String[] args) throws InterruptedException {
8*         Consumer consumer = new Consumer();
9*         Thread t1 = new Thread(consumer);
10*        t1.start();
11*    }
12*
13*    public void run() {
14*        while (true) {
15*            try {
16*                // Create a ConnectionFactory
17*                ActiveMQConnectionFactory connectionFactory = new ActiveMQConnectionFactory(
18*                        "tcp://localhost:61616");
19*
20*                // Create a Connection
21*                Connection connection = connectionFactory.createConnection();
22*                connection.start();
23*            }
24*        }
25*    }
26*
27* }
28*
29* 
```

Figure 5: Consumer maven project

```

1 package ActiveMQProducer.ActiveMQProducer;
2
3 import javax.jms.Connection;
4
5 public class Producer {
6    public static void main(String[] args) {
7        try {
8            // Create a ConnectionFactory
9            ActiveMQConnectionFactory connectionFactory =
10            connectionFactory.createConnection();
11
12            // Create a Session
13            Session session = connection.createSession();
14
15            // Create the destination (Topic or Queue)
16            Destination destination = session.createQueue("TEST");
17
18            // Create a MessageProducer from the Session
19            MessageProducer producer = session.createProducer(destination);
20            producer.setDeliveryMode(DeliveryMode
21            
```

Figure 6: Producer maven project

Producer Class -

```

import javax.jms.Connection;
import javax.jms.DeliveryMode;
import javax.jms.Destination;
import javax.jms.MessageProducer;
import javax.jms.Session;
import javax.jms.TextMessage;
import org.apache.activemq.ActiveMQConnectionFactory;

public class Producer {
    public static void main(String[] args) {
        try {
            // Create a ConnectionFactory
            ActiveMQConnectionFactory connectionFactory = new
            ActiveMQConnectionFactory("tcp://localhost:61616");

            // Create a Connection
            Connection connection = connectionFactory.
            createConnection();
            connection.start();

            // Create a Session
            Session session = connection.createSession(false,
            Session.AUTO_ACKNOWLEDGE);

            // Create the destination (Topic or Queue)
            Destination destination = session.createQueue("TEST");
        
```

Continued to page 67...



This is the third part in this series of articles and is about learning the new features introduced in JavaScript as part of the ECMAScript 2015 specification. Let's look at these new features — computed property keys, String helper functions and additions to Number objects.

In the second part of this series, we learned about binary/octal literals and enhanced object literals. We also looked at different levels of support in JS environments. Since September this year, an important event has happened—the merging of *Node.js* and *io.js*. *Node.js* is the community-led foundation and Joyent backed open source project. *io.js* is the community fork formed with the intention of having a truly community driven JS environment and bringing in ES6 features faster. Both groups have decided to merge and *Node.js* version 4.0 is the resultant combination. For a detailed insight into the significance of the merger, read the article in the first reference given at the end of this article.

Something good about what we are going to learn this month is that these features are supported in most common JS environments, i.e., *Node.js* 4.0+, Babel, Chrome 45 and Firefox 41.

Computed property keys

In JavaScript, an object is used to group variables into a collection. Typically, the creation of an object involves defining a set of properties and assigning default values. In addition, JavaScript has the capability of dynamically (after definition) adding and deleting properties on-the-fly. The property names at the time of definition are strings, which follow rules of definition of a variable and are fixed, as in the example below:

```
let point = {
  x: 10,
  y: 20,
}
```

In ES6, JavaScript comes with the capability of defining property names as computed strings. This capability gives you programmatically defined property names that could change with each run and also adapt property names to the environment. Let us look at the various ways in which the property names can be computed:

```
1. let i = 0;
2. let std = "ecma";
3. let fn = "print";
4.
5. let obj = {
6.   ["js" + ++i]: i,
7.   ["js" + "2"] : ++i,
8.   ["js" + new Date().getFullYear()]: 10,
9.   [std]: 20,
10.  [fn]: function() { console.log("hello!"); },
11. };
12. console.log(JSON.stringify(obj)); // {"js1":1,"js2":2,"js
2015":10,"ecma":20}
13. obj.print(); // hello!
```

In the above code snippet, lines 6 to 10 have different ways in which the property name is computed. In line 6, we have created a property name concatenating a string with a variable value. In line 7, we have concatenated two strings. In line 8, we have part of the property name as the return value of a function. With such capabilities, the program can have property names change at the time of execution. The same program, run in 2016, will give the property name *js2016*. In line 10, we have an example of a function name computed from a variable substitution.

String functions

There are new string functions that will be part of the JavaScript language with effect from the ES6 version. They can be categorised into template functions and search functions, as shown in Table 1.

Table 1

Category	Functions	Type
Template functions	raw``	Static
Search functions	<i>startsWith()</i> <i>endsWith()</i> <i>includes()</i> <i>repeat()</i>	<i>prototype</i> <i>prototype</i> <i>prototype</i> <i>prototype</i>

You can see in Table 1 that new functions are of two types - static and prototype. Static functions are available without an instance variable, whereas a prototype function should have a string instance. This will be clear in the following examples.

The *String.raw* function

This function is introduced along with template strings (refer to Part 1 of this series). Template strings allow variable substitution, and the interpretation of special characters like newline, tab, etc. During string operations, it is required to search or compare characters in the exact raw form instead of interpreted forms. The *String.raw* function keeps the string the same as what is ‘seen’ by the programmer in the editor. In other words, ‘\n’ is two characters ‘\’ and ‘n’.

String.raw is a static function, which means you do not need to have a string object created to use the function.

The raw function has two forms:

```
String.raw`template string`
String.raw({ raw: 'string' }, substitute_string)
```



Note: The first form does not have parenthesis like other functions. It is just a template string enclosed in backticks. In the second form the property name should be ‘raw’.

1. let year = "2015";
2. console.log(`osidays\t\${year}`); //osidays 2015
3. console.log(String.raw`osidays \t\${year}`); //osidays \t2015
4. console.log(String.raw({ raw: 'osi' }, '1', '2')); // 01s2i
5. console.log('osi \uD83D\uDE80'); //osi
6. console.log(String.raw`osi \uD83D\uDE80`); //osidays \uD83D\uDE80

To see the special characters in the above code snippet, please visit, <http://www.es6fiddle.net/ifxcl0ay/>.

Line 2 is an example of a template string. Line 3 displays the non-printable tab character in the raw form. Line 4 is the second form of a raw function where it takes two arguments. The first is the string parameter and the second is a substitution string. Line 5 prints special Unicode characters. In line 6, the same string with *String.raw* displays individual characters.

Finding string and repeat

New helper string functions introduced in ES6 are *startsWith()*, *endsWith()* and *includes()*. These three functions need objects to be created, before they can be used. They also take optional second parameters to indicate from which character position the *find* operation should start.

```
let str = "osidays2015";
console.log(str.startsWith("osi")); // true
console.log(str.startsWith("si", 2)); // false
console.log(str.endsWith("15")); // true
console.log(str.endsWith("20")); // false
console.log(str.includes("day")); // true
```

The next utility function is *repeat()*, which takes an argument and returns the same string ‘count’ a number of times.

```
let pattern = "=+=";
console.log(pattern.repeat(5)); //=====+=+=+=+=+
console.log('ab'.repeat(5)); //abababab
```

Number object extensions

The already existing global functions, *isFinite()*, *isNaN()*, *parseFloat()* and *parseInt()*, are now added to the ‘Number’ object. This groups related functions under the *Number* object. One difference we should be aware of is that the *global* function automatically converts a number in the string form to a number, whereas when used with the *Number* object, it does not. The following example makes this clear:

```
console.log(Number.isFinite(3)); // true
console.log(Number.isFinite('3')); // false
console.log(isFinite('3')); // true
```

In addition to functions, there are three properties added to the *Number* object. They are: *EPSILON*, *MIN_SAFE_INTEGER* and *MAX_SAFE_INTEGER*.

JavaScript stores numbers as floating point numbers internally. Arithmetic operations may not yield the desired results due to a difference in small precision. For example, $0.21+0.2 \neq 0.41$ due to precision error.

The *Number.EPSILON* is the smallest interval between two representable numbers. Its value is $2.220446049250313e-16$.

To overcome precision errors, *Number.EPSILON* is used. When a math operation is performed, check if the resultant difference is more than *Number.EPSILON*. As we are aware, JavaScript has number representation that is in the range of 2^{-53} and 2^{+53} .

```
console.log(Number.MIN_SAFE_INTEGER); // -9007199254740991
console.log(Number.MAX_SAFE_INTEGER); // 9007199254740991
```

The new properties help in checking the valid range after mathematical computations. There is a function *isSafeInteger()* that checks if a number is within valid range.

```
console.log(Number.isSafeInteger(42342343434343)); // true
```

```
console.log(Number.isSafeInteger(23234234234343434)); //  
false
```

Installing Node.js 4.1.1 in Ubuntu

It is important if you have migrated to the latest *Node.js* environment to try ES6 features. The following two steps will get you to the latest version. If yours is not an Ubuntu environment, please refer to the *Node.js* home page.

```
curl -sL https://deb.nodesource.com/setup_4.x | sudo -E bash -  
sudo apt-get install -y nodejs
```

Staying up-to-date

If you are interested in learning more about JavaScript, take a look at Table 2, which features a few influential people you should follow.

Next month we will learn about the *for..of* iterator which

Table 2

Twitter handle	Name	Known for
@rauschma	Axel Rauschmayer	Rauschmayer is the author of the book, 'ExploringJS' and blogs at @2ality
@BrendanEich	Brendan Eich	Invented JavaScript and co-founded Mozilla
@paul_irish	Paul Irish	Core skills in Chrome DevTools and performance
@addyosmani	Addy Osmani	Engineer at Google working on Chrome and Polymer

can be used on strings and collection data structures like arrays. We will also learn about a shorthand notation called arrow functions.

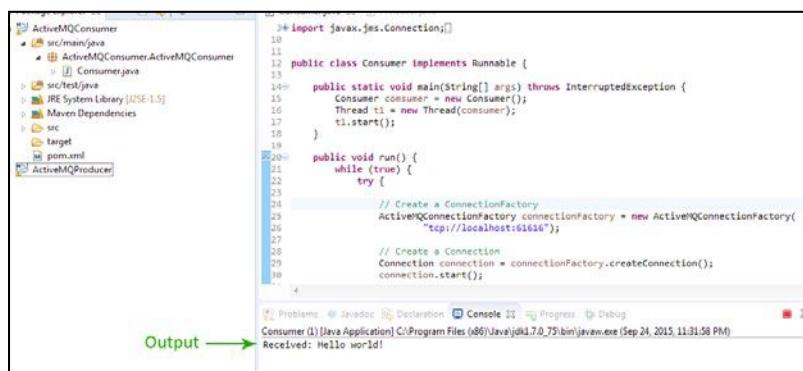
References

- [1] 'io.js and Node.js Have United and That's a Good Thing'-- <http://thenewstack.io/io-js-and-node-js-have-united-and-thats-a-good-thing/>
- [2] Detailed list of JS environments' support: <http://kangax.github.io/compat-table/es6/>
- [3] Try ES6 features online <http://www.es6fiddle.net/>

By: Janardan Revuru

The author works at HP EG India R&D, Bengaluru. He is passionate about JavaScript and related Web frameworks. He can be reached at janardan.revuru@gmail.com

Continued from page 64...



```
ActiveMQConsumer  
src/main/java  
  ActiveMQConsumer.ActiveMQConsumer  
    Consumer.java  
src/test/java  
  JRE System Library [JSE-1.5]  
  Maven Dependencies  
src  
target  
pom.xml  
  
ActiveMQProducer  
  
2+ import javax.jms.Connection;  
3+  
4+ public class Consumer implements Runnable {  
5+  
6+   public static void main(String[] args) throws InterruptedException {  
7+     Consumer consumer = new Consumer();  
8+     Thread t1 = new Thread(consumer);  
9+     t1.start();  
10+  }  
11+  
12+  public void run() {  
13+    while (true) {  
14+      try {  
15+        // Create a ConnectionFactory  
16+        ActiveMQConnectionFactory connectionFactory = new ActiveMQConnectionFactory(  
17+          "tcp://localhost:61616");  
18+  
19+        // Create a Connection  
20+        Connection connection = connectionFactory.createConnection();  
21+        connection.start();  
22+  
23+        // Create a Session  
24+        Session session = connection.createSession(false,  
25+          Session.AUTO_ACKNOWLEDGE);  
26+  
27+        // Create a MessageProducer from the Session to  
28+        // the Topic or Queue  
29+        MessageProducer producer = session.  
30+        createProducer(destination);  
31+        producer.setDeliveryMode(DeliveryMode.PERSISTENT);  
32+  
33+        // Create a messages  
34+        String text = "Hello world!";  
35+        TextMessage message = session.  
36+        createTextMessage(text);  
37+        producer.send(message);  
38+      } catch (Exception e) {  
39+        System.out.println("Caught: " + e);  
40+        e.printStackTrace();  
41+      }  
42+    }  
43+  }  
44+}  
  
Problems Declaration Console Progress Debug  
Consumer (1) Java Application C:\Program Files (x86)\Java\jdk1.7.0_75\bin\java.exe (Sep 24, 2015, 11:31:58 PM)  
Received: Hello world!
```

Figure 7: Consumer consuming the message

```
Destination destination = session.createQueue("TEST");  
  
// Create a MessageProducer from the Session to  
the Topic or Queue  
MessageProducer producer = session.  
createProducer(destination);  
producer.setDeliveryMode(DeliveryMode.PERSISTENT);  
  
// Create a messages  
String text = "Hello world!";  
TextMessage message = session.  
createTextMessage(text);  
producer.send(message);
```

```
// Clean up  
session.close();  
connection.close();  
} catch (Exception e) {  
  System.out.println("Caught: " + e);  
  e.printStackTrace();  
}  
}
```

In the above example, the delivery mode is set to PERSISTENT, which means that even if the ActiveMQ server goes down, the *Producer* data persists.

Now, run the *Consumer* application. The *Consumer* will keep checking whether there is any message to consume from ActiveMQ.

Next, run the *Producer* application to send messages to the ActiveMQ. Each run of the *Producer* will send a 'Hello World!' message to ActiveMQ and, in turn, the *Consumer* will consume the message from ActiveMQ.

By: Sumeet Patil

The author is a middleware expert, open source enthusiast and likes to explore new technologies. He has worked on ActiveMQ, Servicemix, Apache Camel, SAP(ABAP, SAP UI5, Gateway etc), Java/J2EE, Android native/ hybrid, HTML5, NodeJs, PHP and more. He can be reached at sumeetp1991@gmail.com



The Importance of Memory Management

Memory management requires that the programmer provides ways to dynamically allocate portions of memory to programs, when requested, and free it for reuse when it is no longer needed. In any advanced computer system, where more than a single process might be running at any given point in time, this is critical.

When any program is loaded into the memory, it is organised into three parts called segments, which are:

- Text segment (the code segment)
- Stack segment
- Heap segment

Text segment

The text segment is where the compiled code of the program itself resides. This is the machine language representation of the program steps to be carried out, including all functions making up the program, both user and system defined. So, in general, an executable program generated by a compiler (like GCC) will have memory organisation, as shown in Figure 1. This segment is further categorised as follows.

Code segment: This contains the code (executable or code binary).

- Data segment:** This is sub-divided into two parts.
- *Initialised data segment:* All the global, static and constant data are stored in the data segment.
 - *Uninitialised data segment:* All the uninitialised data are stored in the Block Started by Symbol (BSS).

Stack segment

The stack is used to store our local variables, and for passing arguments to the functions along with the return address of the instruction that is to be executed after the function call is over. When a new stack frame needs to be added (as a result of a newly called function), the stack grows downwards.

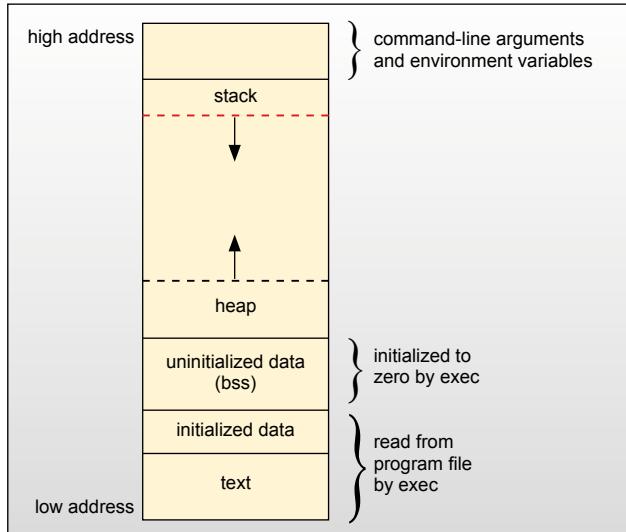


Figure 1: Memory organisation of a typical program code

Heap segment

When the program allocates memory at runtime using the `calloc` and `malloc` functions, then memory gets allocated in heap. When some more memory needs to be allocated using the `calloc` and `malloc` functions, heap grows upwards.

Memory management in C

There are two ways in which memory can be allocated in C:

- By declaring variables
- By explicitly requesting space from C

The C programming language provides several functions for memory allocation and management. These functions can be found in the `<stdlib.h>` header file.

All the functions are given below:

- `void *calloc(int num, int size)`: This function allocates an array of `num` elements. The size of each element in bytes will be `size`.
- `void free(void *address)`: This function releases a memory block specified by an address.
- `void *malloc(int num)`: This function allocates an array of `num` bytes and leaves them initialised.
- `void *realloc(void *address, int newsize)`: This function reallocates memory extending it up to `newsize`.

Allocating memory dynamically

While programming, if we are aware of the size of an array, it is easy to define the array. For example, to store the name of any person, if the array can go up to a maximum of 100 characters we can define something as follows:

```
char name[100];
```

But now, let us consider a situation where we have no idea about the length of the text we need to store. For example, if we want to store a detailed description about a topic, we need

to define a pointer to a character without defining how much memory is required and later, based on the requirement, we can allocate memory, as shown in the example below:

```
#include <stdio.h>
#include <stdlib.h>
#include <string.h>
int main()
{
    char name[100];
    char * description;
    strcpy(name, "Vikas Awasthy");
    /* allocate memory dynamically */
    description = malloc( 200 * sizeof(char) );
    if( description == NULL )
    {
        fprintf(stderr, "Error - unable to allocate required memory
\n");
    }
}
```

```
else
{
    strcpy( description, "Vikas Awasthy is an Application
Developer");
}
printf("Name = %s\n", name );
printf("Description: %s\n", description );
}
```

Output of Code:-

```
Name = Vikas Awasthy
Description: Vikas Awasthy is an Application Developer
```

Resizing and releasing memory

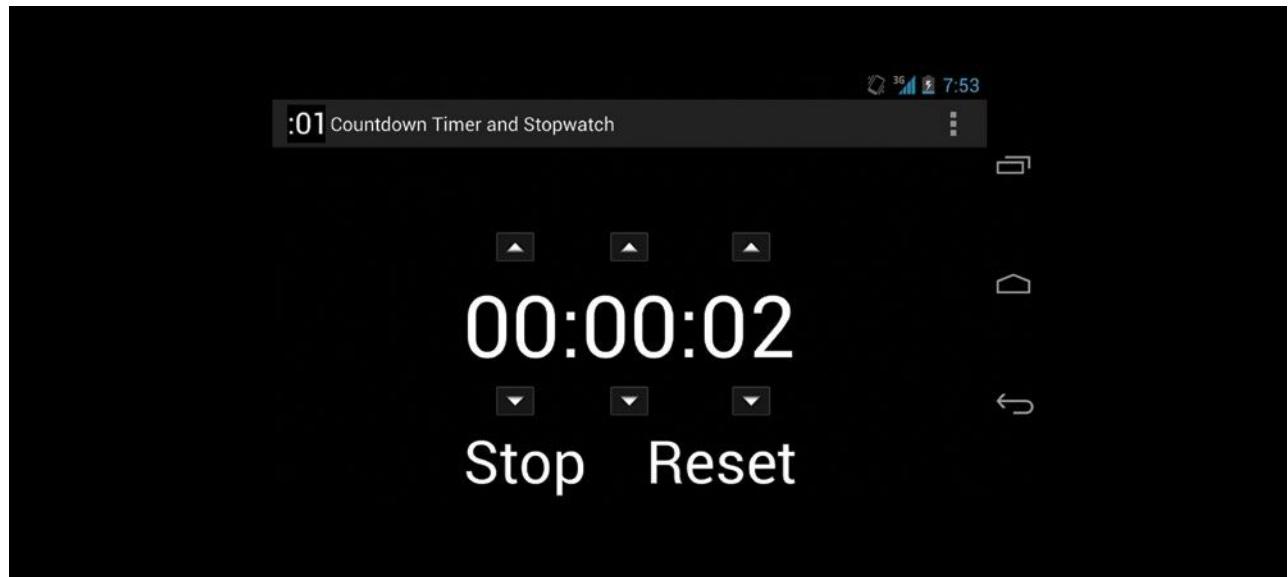
When our program is executed, the operating system automatically releases all the memory allocated by our program but as a good practice, when we are not in need of memory any more, then we should release that memory by calling the function `free()`.

Alternatively, we can increase or decrease the size of an allocated memory block by calling the function `realloc()`. Let us check the above program once again and make use of the `realloc()` and `free()` functions:

```
#include <stdio.h>
#include <stdlib.h>
#include <string.h>
int main()
{
    char name[100];
    char * description;
```

Building Manual and Automatic Counters in App Inventor 2

As this series has progressed, some readers who have been following it may have built a few apps or even a collection of them. This article guides the reader through the steps involved in building counters – both manual and automatic.



As we proceed further in making Android apps using the open source App Inventor tool, let's now introduce a little of mathematics into the application. The next project will be a manual counter, before we go on to building an automatic counter run by the internal clock of the phone, which will be handier.

Manual counter

Though mathematics is not liked by some of us, its role in our daily lives can't be ignored. Every single thing that requires keeping track of has to go through a counting process where the use of the most basic mathematics is required. We start learning mathematics via numbers and counting, so here as well, we will begin with a number game to count things. We are going to create a manual counter application, which, on the press of a button, enables you to increase the count and on the press of another button enables you to decrease the count.

GUI requirements

We will not need much by way of GUI requirements as the app is very basic. We will require two button components and a label component, along with some horizontal arrangements to keep things aligned.

We will require the following components for this project:

	Component	Purpose	Location
1	Horizontal arrangement	To arrange the child components horizontally	Palette--> Layout--> Horizontal Arrangement
2	Label	To display the value of what has been counted	Palette-->User Interface-->Label
3	Button	One button to increase the count	Palette-->User Interface-->Button
4	Button	One button to decrease the count	Palette-->User Interface-->Button

Getting started

1. Drag and drop the components mentioned above to the viewer.
2. The visible components can be seen by you while the non-visible components will be located beneath the viewer under the tag, 'Non-visible'.
3. We have placed a label so as to put the name of the application.



Figure 1: Manual counter designer screen

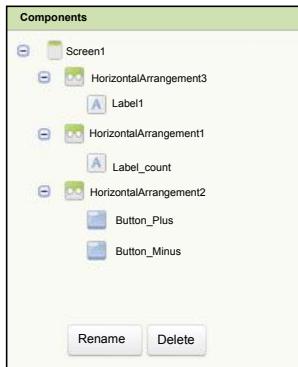


Figure 2: Manual counter components

4. Both buttons have been put within the ‘Horizontal arrangement’, so they are aligned horizontally.
5. We have also changed the background colour property of both buttons to identify them more clearly.
6. If you have dragged and placed everything, the layout will look somewhat as in Figure 1.
The components viewer will look like what’s shown in Figure 2.
7. Make the necessary property changes like we did earlier when changing the text property, for the label and button components.
8. So this way, your graphical user interface is ready. Figure 1 shows exactly what the application will look like after installation.

Before heading towards the blocks editor to define the behaviour, let’s discuss what actual functionality we expect from our application.

1. When we press the ‘Plus’ button, the counter value should increase.
2. When we press the ‘Minus’ button, the counter value should decrease.

So let’s move on and add these two features using the block editor. I hope you remember how to switch from the designer to the block editor. There is a button available right above the ‘Properties’ pane to switch between the two.

Block editor blocks

1. The blocks shown in Figure 3 are easy to understand. These indicate that when the button ‘Plus’ is pressed, it should increase whatever figure is shown as the count label.
2. When the button ‘Minus’ is pressed, it should decrease whatever figure is shown as the count label.
3. The additional ‘Plus’ and ‘Minus’ blocks you see fall under the mathematics category.

The automatic counter

In the last section, you counted things manually using buttons. You pressed the ‘Plus’ or ‘Minus’ buttons to get the desired result. You will now be able to build another useful app,

which is a simple countdown timer that will be triggered on pressing a button and will show you the time elapsed in seconds, minutes or hours.

GUI requirements

Not much is needed by way of GUI requirements as the whole thing is very basic. We will require two button components and a label component, along with some horizontal arrangements to keep things aligned.

We will require the following components for this project:

	Component name	Purpose	Location
1	Horizontal arrangement	To arrange the child components horizontally	Palette-->Layout-->Horizontal Arrangement
2	Label	To display the count value	Palette-->User Interface-->Label
3	Button	One button to increase the count	Palette-->User Interface-->Button
4	Button	One button to decrease the count	Palette-->User Interface-->Button
5	Clock	To count time	Palette-->Sensors-->Clock

1. Drag and drop the components mentioned above to the viewer.
2. Visible components can be seen, while the non-visible components will be located beneath the viewer under the tag ‘Non-visible’.
3. We have placed a label to put the name of the application.
4. All buttons have been put within the ‘Horizontal arrangement’ so as to keep them aligned horizontally.
5. We have also changed the background colour property of both buttons to identify them more easily.
6. You will see the clock component for the very first time. It has access to the inbuilt clock of the phone and performs actions accordingly.
7. If you have dragged and placed everything, the layout will look something like what’s shown in Figure 4.

Figure 5 shows what the components viewer will look like. Since you are not very familiar with the clock component and its properties, set the properties as depicted in the picture shown in Figure 6.

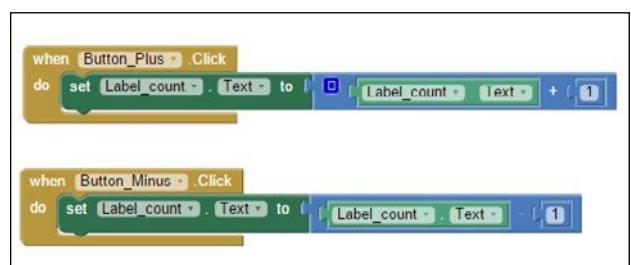


Figure 3: Manual counter block editor blocks

8. Make the necessary property changes like we did in changing the text property, for the label and button components.
9. Now your graphical user interface is ready. Figure 4 shows exactly what the application will look like after the installation.

Now let's head towards the block editor to define the features, so that we can determine the actual functionality that we expect from our application.

1. When we press the 'Start' button, the counter value should increase and continue to do so.
2. When we press the 'Stop' button, the counter value should be paused to the current value.
3. When we press the 'Reset' button, the counter value should be reset to the default 0.

So let's move on and add these two features using the block editor.

Block editor blocks

1. The blocks shown in Figure 7 are easy to understand. These say that when the button 'Start' is pressed, it should increase whatever figure is visible as the count label.
2. When the button 'Stop' is pressed, it should stop the timer at the current value. For stopping the clock, we do not need to set it to 'disable'; instead, we need to set 'enable' as false.
3. When the 'Reset' button is pressed, it stops the clock and sets the counter value to the default 0.
4. The fourth block is the one that you will see for the very first time. It indicates that when the timer is enabled, it will increase the label value by one in each repeat cycle. If you remember the properties of the clock component, we have set the timer interval as 1000 milliseconds, which means the clock will have a new cycle every second and hence will be increasing the value.

Now you are done with the block editor too. Next, we will move on to download and install the app on your phone to check how it works.

Packaging and testing

To test the app, you need to get it on your phone. First, you have to download the application to your computer and then move it to your phone via Bluetooth or USB cable. I'll tell you how to download it.

1. On the top row, click on the 'Build' button. It will show you the option to download the apk to your computer.
2. As the downloading is going on, the progress can be seen and after a successful download, the application will be placed in the download folder of your directory or in the location you have set for it.
3. Now you need to get this apk file to your mobile phone either via Bluetooth or USB cable. Once you have placed the apk file on your SD card, you need to install it. Follow the on-screen instructions to install it. You might get some notification or warning saying, "Installing from un-trusted



Figure 4: Automatic counter designer screen

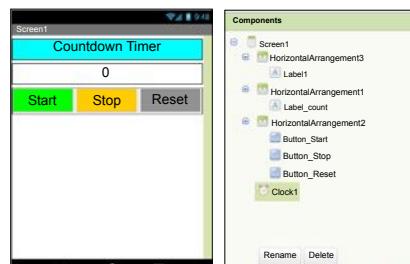


Figure 5: Automatic counter components

Properties	
Clock 1	
TimerAlwaysFiles	<input type="checkbox"/>
TimerEnabled	<input type="checkbox"/>
TimerInterval	1000

Figure 6: Properties for clock component



Figure 7: Automatic counter block editor blocks

source." Allow this from the *Settings* and after successful installation, you will see the icon of your application in the menu of your mobile. Here, you will see the default icon, which can be changed (we will tell you how to do so as we move ahead).

I hope your application is working exactly as per your requirements. Now, depending upon your usability and customisation, you can change various things—buttons, text and behaviour too.

Debugging the application

As of now, we have just created the prototype of the application with very basic functionality. What else would a user be interested in? Take care that certain features of the app do not annoy the user and check if your app is able to sustain itself. Consider the following cases:

1. Don't you think there should be a reset button to clear the value and start counting from the beginning?
2. If you were asked to add an additional reset button, will you be able to do so?
3. What extra GUI feature do you think should be added?
4. How many blocks should be added to furnish the reset functionality?
5. Will it be useful to store the counter value in the database and retrieve it if needed?

These are some of the scenarios that might occur and users will be pretty happy seeing them implemented.

Think over all the needs that may crop up, and how you could integrate these into the application. Do revert me if you fail to accomplish any of the above features. 

By: Meghraj Singh Beniwal

The author is a B. Tech in electronics and communication, a freelance writer and an Android apps developer. He is currently working as a systems engineer at Infosys, Mysore. He can be contacted at meghrajsingh01@rediffmail.com.

Creating Apps on Live Cloud Platforms: IBM Bluemix and Red Hat OpenShift

Cloud platforms are becoming rather popular with developers for various well known reasons. Through the use of virtual machines, developers can avail of the various cloud services providers offer. This article serves as a primer on creating apps on IBM Bluemix and Red Hat OpenShift.

Nowadays, most computing services that are hosted by a number of cloud service providers are available on demand. Currently, a number of key areas in cloud computing are being researched due to the escalating needs in the digital services domain.

Virtualisation and hypervisors

Virtualisation is the key technology that works at the back end of cloud services and digital infrastructure. In actual implementation, the cloud is deployed using virtualisation. Whenever there is a need for a remote system by any user, dynamically created virtual machines are provided to the end user or developer at the other remote location with the credentials of the machine at the cloud data centre.

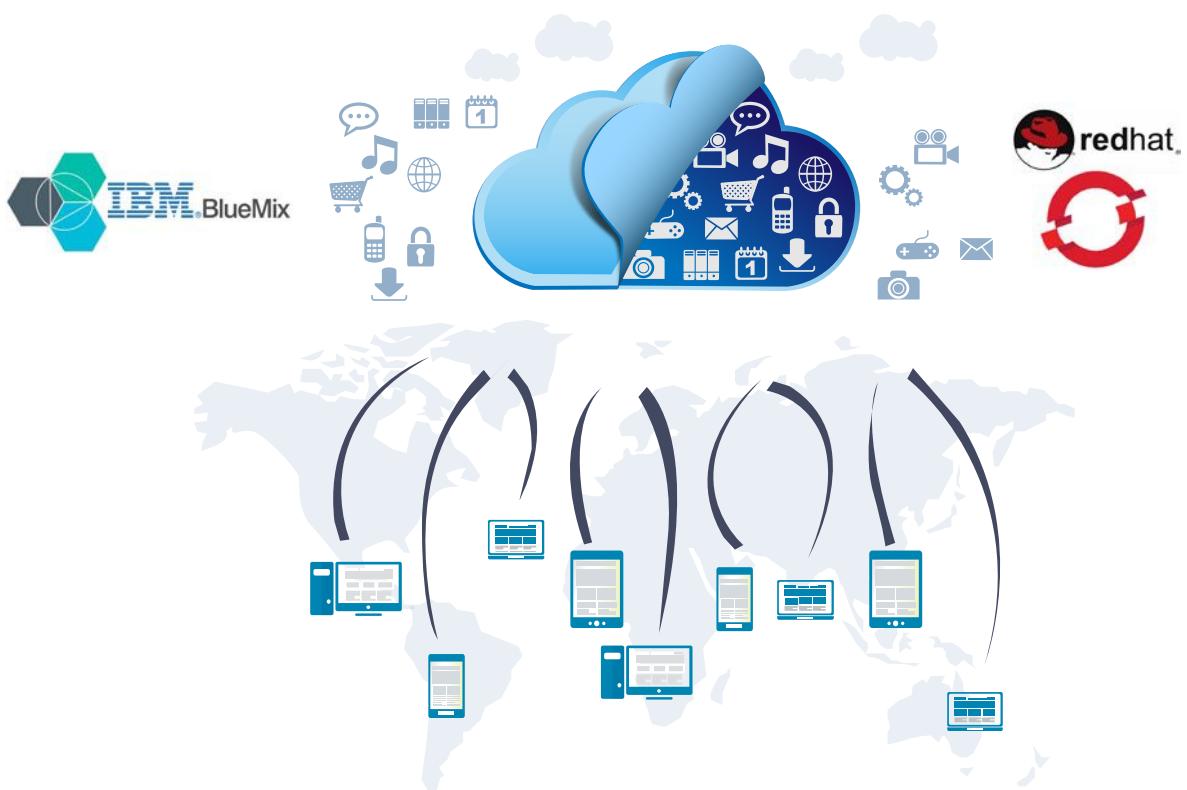
A virtual machine (VM) refers to the software implementation of any computing device, machine or computer which executes the instructions or programs exactly as the physical machine would. When a user or developer works on a

virtual machine, the resources, including all programs installed on the remote machine, are accessible using a specific set of protocols. Here, for the end user of the cloud service, the virtual machine acts like the actual machine.

Widely used virtualisation software includes VMware Workstation, VMlite Workstation, Virtual Box, Hyper-V, Microsoft Virtual PC, Xen and KVM (Kernel Based Virtual Machine).

Hypervisor technology

A hypervisor or virtual machine monitor (VMM) is a piece of computer software, firmware or hardware that creates and runs virtual machines. A computer on which a hypervisor is running one or more virtual machines is defined as a host machine. Each virtual machine is called a guest machine. The hypervisor presents the guest operating systems with a virtual operating platform and manages the execution of these systems. Multiple instances of a variety of operating systems may share the virtualised hardware resources.



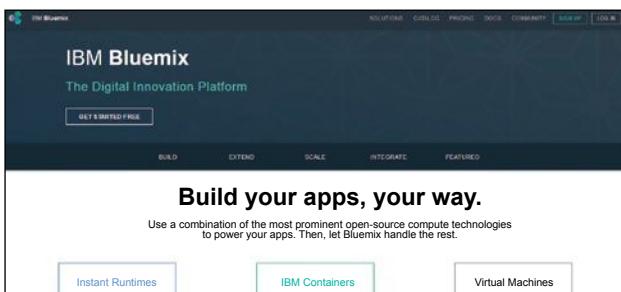


Figure 1: The IBM Bluemix online platform

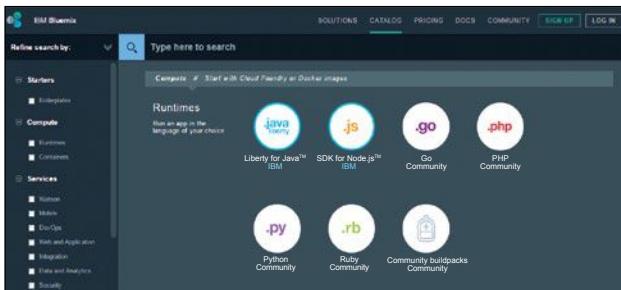


Figure 2: IBM Bluemix in a Platform as a Service (PaaS) mode

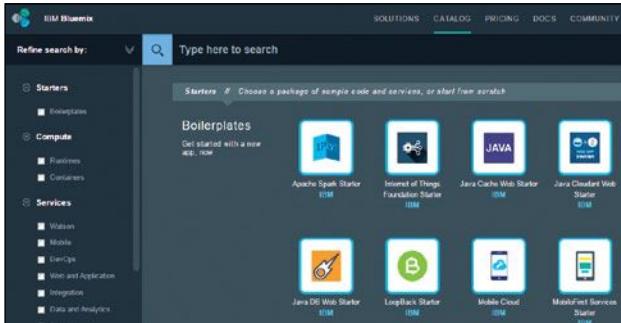


Figure 3: Selecting the IoT application from IBM Bluemix

Table 1: Hypervisors used in the industry

Hypervisor	Cloud service provider
Xen	<ul style="list-style-type: none"> Amazon EC2 IBM SoftLayer Fujitsu Global Cloud Platform Linode OrionVM
ESXi	VMware Cloud
KVM	<ul style="list-style-type: none"> Red Hat HP Del Rackspace
Hyper-V	Microsoft Azure

IBM Bluemix - *bluemix.net*

IBM Bluemix is a leading cloud service provider offering a number of services including Big Data analytics, mobile applications, NoSQL databases, the Internet of Things (IoT), natural language processing, sentiment analysis, and a lot of related high performance computing services.

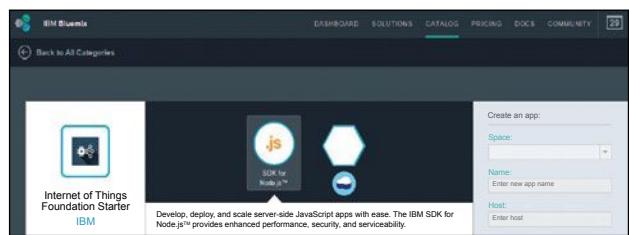


Figure 4: Specifying the app name in IBM Bluemix

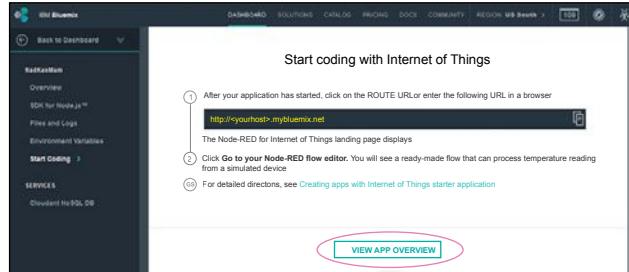


Figure 5: View the app details

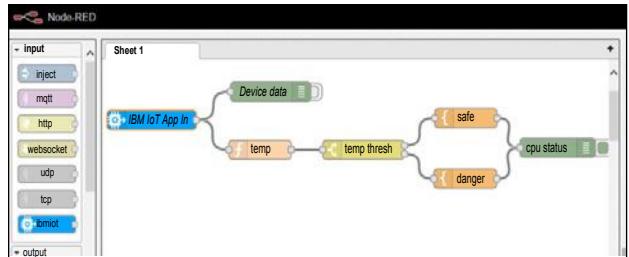


Figure 6: Configuring the IoT app in Node RED

The IBM Bluemix Platform can be accessed using the ‘freemium’ model, and there is the provision to use boilerplates. The latter simplifies the creation and updation of live code, enabling even beginners to use IBM cloud services without any difficulty.

In IBM Bluemix, there is a detailed catalogue of a number of cloud services including the programming platform for PHP, Java, Ruby, Python and many other languages. Bluemix can be used as a *PaaS* for hosting and running multiple applications.

After logging in, click on the *Catalog* button to view and select different platforms and services provided by IBM Bluemix.

Creating an Internet of Things (IoT) application in Bluemix

Let’s create an IoT application in IBM Bluemix. On the *Catalog* page, click ‘Internet of Things Application Foundation Starter’. This is a simulation scenario for sensing the temperature.

As in Figure 4, the app name is to be specified while creating the live cloud application. Using this unique name, the application is identified and executed on the cloud servers.

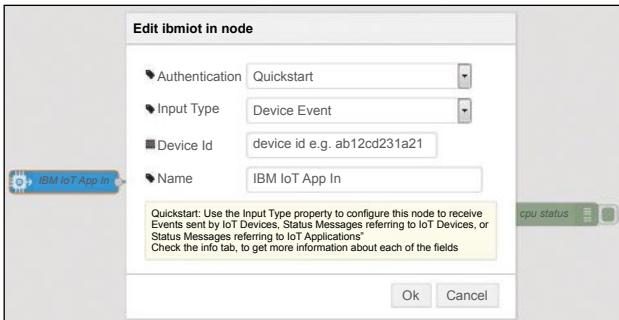


Figure 7: Setting up the IoT sensor

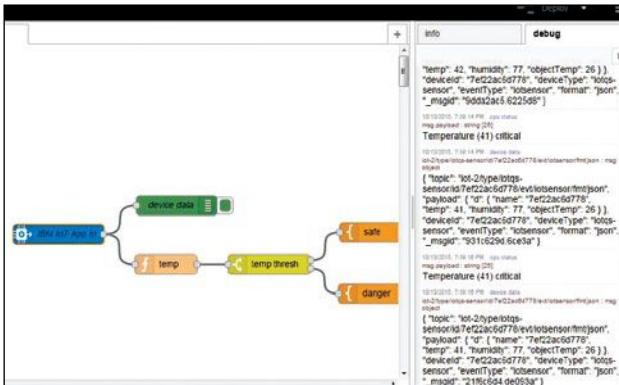


Figure 8: Analysing and logging the status of the sensor on changing the temperature



Figure 9: OpenShift login panel

You can get a view of the temperature sensor simulator at <https://quickstart.internetofthings.ibmcloud.com/iotsensor/>.

In Figure 7, the device ID is mapped with the actual device. Let's assume there is a particular machine and we want to sense its temperature using IoT. In this case, the unique machine ID is mentioned in Node RED so that the application can sense the live temperature or any other required parameter.

In Figure 8, the live logging and analysis of the temperature can be viewed. Suppose the machine under analysis is moved beyond the threshold temperature, the IoT application will sense and display the appropriate message mentioned in the code.

Creating PHP applications in IBM Bluemix

IBM Bluemix provides a PHP starter application template to

The screenshot shows the 'Welcome to OpenShift' page. It provides instructions: 1. Choose a web framework or codebase to start from (Try Java, PHP, Python, Ruby, Node.js, or create a new Drupal or WordPress site instantly). 2. Add cartridges like MySQL or MongoDB to your application (OpenShift lets you add services and tools to your application through cartridges - including databases, cache servers, management tools, and continuous integration servers). 3. Upload your code to OpenShift via Git (Your source code is stored with your application in a Git version control repository). A blue 'Create your first application now' button is visible.

Figure 10: Creating the first application panel in the OpenShift platform

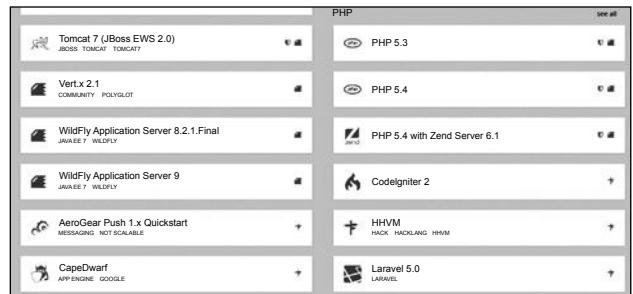


Figure 11: Selecting the PHP platform in OpenShift temperature

The screenshot shows the configuration page for a PHP application. It includes sections for 'Based On' (PHP 5.3 Cartridge), 'Public URL' (http://php.magmaambala.rhcloud.com), 'Source Code' (Optional URL to a Git repository), and 'Gears' (small). A note states: 'Gears are the application containers running your code. For most applications, the small gear size provides plenty of resources. You can also upgrade your plan to get access to more gear sizes.'

Figure 12: Setting up the URL and related parameters of the app in OpenShift

cloud users so that their own code can be added and pushed back to the cloud server. This starter application template is the boilerplate for PHP Web application development in Bluemix.

Initially, a new PHP starter application is activated and starter code is downloaded. After that, the App name and host are provided by the user. For further implementation, the *cf* (Cloud Foundry) command line tool is used. *cf* is the CLI (command line interface) that can be used to deploy and control the applications on Bluemix.

To log in to the Bluemix environment, type:

```
$ cf login -a https://api.ng.bluemix.net -o <myorg name> -s <myspacename>
```

After making changes to the code, deploy the application again. For deploying the modified application to Bluemix, the *cf* command is used:

```
$ cf push myappname -p <PathToApplication> -m  
512M
```

Accessing the application

Once the application is deployed, open the URL <http://myappname.mybluemix.net> in a browser to check that it is running.

Creating a PHP application in Red Hat OpenShift

OpenShift (<http://www.openshift.com>) has gained a lot of popularity with developers as it has enormous speed, which makes the quick spinning up of a new development stack possible.

To start using the OpenShift platform, the very first step a developer should take is to open an account that provides access to create and deploy applications.

After creating the new account, there is the option to create a new application. In the online platform, the complete code can also be uploaded using Git so that the pre-built application can be launched on the OpenShift platform.

To work on a new application, the OpenShift platform presents a number of options which the developer may use. If a Python application is to be developed, OpenShift offers an online platform to program and execute Python code.

After the basic settings and selection of the public URL, a new application is created on which the code can be deployed and executed.

The changes in the existing code can be made using

Continued from page 69...

```
strcpy(name, "Vikas Awasthy");
/* allocate memory dynamically */
description = malloc( 30 * sizeof(char) );
if( description == NULL )
{
    fprintf(stderr, "Error - unable to allocate required
memory\n");
}
else
{
    strcpy( description, "Vikas Awasthy is an Application
Developer.");
}
/* suppose we want to store more description */
description = realloc( description, 100 * sizeof(char) );
if( description == NULL )
{
    fprintf(stderr, "Error - unable to allocate required
memory\n");
```

Your application has been created. Continue to the application overview page.

Making code changes

Install the Git client for your operating system, and from your command line run

```
git clone ssh://566a2c8889fcff799400000@myphp-
magmaambala.rhcloud.com/~git/myphp.git/
cd myphp/
```

This will create a folder with the source code of your application. After making a change, **add**, **commit**, and **push** your changes.

```
git add .
git commit -m 'My changes'
git push
```

When you push changes the OpenShift server will report back its status on deploying your code. The sever will run any of your configured deploy hooks and then restart the application.

Manage your app

The console is convenient, but if you need deeper control try our other client tools

Command-Line

All of the capabilities of OpenShift are exposed through our command line tool, **rhc**. Follow these steps to install the client on Linux, Mac OS X, or Windows.

After installing the RHC, read more on how to manage your application from the command line in our User Guide.

JBoss Developer Studio

The JBoss Developer Studio is a full featured IDE with OpenShift integration built in. It gives you the ability to create, edit and deploy applications without having to leave the IDE. Links to download, install and use the JBoss Developer Studio for Linux, Mac OS X, or Windows can be found on the JBoss Developer Studio tools page.

Figure 13: Git instructions for app creation and upation in OpenShift



Figure 14: Deleting the app from the OpenShift platform

Git. For this, the Git tool is required to be installed so that the modifications and uploading can be done.

In OpenShift, the creation and deletion of an application is very easy. The application created earlier can be deleted any time, after use. 

By: Dr Gaurav Kumar

The author is associated with a number of academic institutes, where he delivers lectures and conducts technical workshops on the latest technologies and tools. You can contact him at kumargaurav.in@gmail.com; URL: www.gauravkumarindia.com.

```
}
```

```
else
```

```
{
```

```
strcat( description, "He works in IBM India.");
```

```
}
```

```
printf("Name = %s\n", name );
```

```
printf("Description: %s\n", description );
```

```
/* release memory using free() function */
```

```
free(description);
```

```
}
```

Output of above code:-

Name = Vikas Awasthy

Description: Vikas Awasthy is an Application Developer. He works in IBM India.



By: Vikas Awasthy

The author is an application developer (Java) and currently works in IBM India Pvt Ltd, Gurgaon.

Haxe: Write Once, Convert to Many Languages

The proliferation of a wide variety of platforms has become a major challenge for developers when it comes to porting their application to different environments. Haxe assists developers to face this challenge. This article is an eye opener to the world of cross-compilation based programming using the Haxe toolkit.



One of the major challenges for software developers today is dealing with the variety of software environments or platforms. Applications can enjoy the fullest potential of the platforms only when they are built using native codebase. To harness this advantage, organisations need to have developers who are skilled in those individual development techniques. Typically, teams address this with the help of coders fluent in specific programming languages to port an application to a specific platform. Now, instead of this traditional approach, there are special multi-platform or multi-paradigm cross-compiler based languages.

Multi-platform languages

Here, the word ‘multi-platform’ is not used to imply that the language will work on different platforms, but rather that the source code of these languages can be compiled into the source code of other programming languages. This can also be called ‘source-to-source compilation’. Two such popular languages are Haxe and Monkey. The code written in these programming languages is compiled into specific target languages, which can again be handled as usual programs written in those target programming languages. Due to this source-to-source translation methodology, all that a

developer needs to do is to write the Haxe code and convert it into the target language, so that the platform-specific features can be harnessed fully.

The Haxe toolkit

Haxe is a popular multi-platform toolkit that has five major components:

- The Haxe programming language
- A cross-compiler
- The Haxe library
- Haxe based frameworks
- Additional tools

Haxe is a high-level programming language. With respect to data types, Haxe is strictly a typed language. The effort required to learn the syntax of Haxe is very minimal if you already know any of the modern day programming languages like Java, PHP, etc.

Why Haxe?

Though there are other alternatives to Haxe in the multi-platform programming world, Haxe possesses the following advantages:

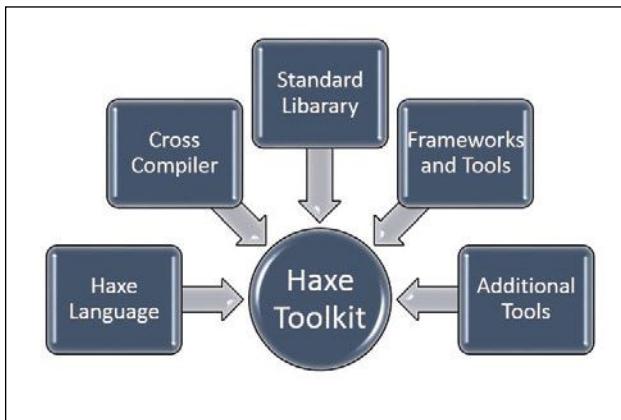


Figure 1: Haxe components

- Haxe is available for download with an open source licence.
- It can be configured in various operating systems without any major glitches.
- The Haxe programming language is fully featured.
- It has the capability to target both programming languages and virtual machines like Neko.
- Haxe has support for modern programming languages like C#, Python, etc.
- It can be used to compile into client side scripts like JavaScript and server side scripts like PHP.

The evolution of Haxe

Development on the Haxe project was started in 2005 by Nicolas Cannasse. Initially, it was named as 'haXe'. The milestones in Haxe's development are listed below:

- The first beta version was released in February 2006. The targets it supported then were AVM bytecode and Neko, which was the virtual machine developed by Cannasse himself.
- The version 1.0 of Haxe was released during May 2006. It had support for JavaScript.
- In August 2006, the 'haxelib' tool was added as a component to the Haxe Toolkit.
- ActionScript 3 was added as a target in March 2007.
- In July 2008, Haxe version 2.0 was released with support for PHP.
- In July 2009, C++ support was added to Haxe with the release 2.04.
- In January 2011, version 2.07 was launched, which supported 'macros'.
- The versions 2.08 and 2.09 focused on the improvements in the JavaScript target.
- In 2012, support for Java and C# was added in Haxe 2.10.
- Haxe 3 was launched in May 2013 with the establishment of the Haxe Foundation to deal with future

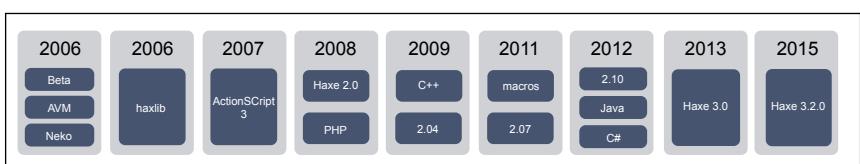


Figure 2: The evolution of Haxe

developments and maintenance.

- The latest version of Haxe is 3.20, which was released this year.

Haxe installation

The latest Haxe version can be downloaded from <http://haxe.org/download/> which provides both the installer and binaries for Windows, Linux and Mac based computers. A typical Haxe installation on a Windows PC is easy (similar is the case with other OSs like Linux). Just download the executable installer file and run it using a double click. Being a small file, the installation will get completed in a few minutes. During the installation, you will be given an option to select or remove the Neko virtual machine too.

The Haxe cross-compiler

The Haxe compiler is capable of compiling the Haxe source code into the source code of the target language. The available programming language targets for the Haxe compiler are given in Table 1.

Language	Potential applications	Available since
JavaScript	Web, desktop, API	Beta (2006)
ActionScript 3	Games, mobile, API	1.12 (2007)
PHP	Web	2.0 (2008)
C++	Games, CLI, mobile, desktop	2.04 (2009)
Java	CLI, mobile, desktop	2.10 (2012)
C#	Mobile, desktop	2.10 (2012)
Python	CLI, Web, desktop	3.2 (2015)

Table 1: Haxe's programming language targets

Apart from the above specified source-to-source compilation, Haxe supports source-to-bytecode compilation as well, as shown in Table 2.

Name	Potential application types	Available since
Flash	Games, mobile	Alpha (2005)
Neko	Web, CLI	Alpha (2005)

Table 2: Haxe virtual machine targets

It is interesting to know that Neko has both programming and virtual machine components.

The Haxe 'Hello World'

The syntax of the Haxe programming language would look very familiar to most programmers as it is more oriented

towards ECMAScript Standards (the scripting language specification standardised by ECMA International in ECMA-262 and ISO/IEC 16262). The detailed documentation on Haxe programming is available for download on <https://github.com/HaxeFoundation/HaxeManual/raw/master/HaxeManual/HaxeManual.pdf>. The Haxe source code can be entered using any text editor. Once the program is entered, it can be compiled into the available target programming languages/bytecode using the Haxe cross-compiler. A sample Hello World program in Haxe is as shown below:

```
class HelloWorld {
    static public function main() {
        trace("Welcome to Haxe Programming");
    }
}
```

This code should be saved as *HelloWorld.hx*. Then, from the command prompt it will be cross-compiled into JavaScript using the following command:

```
haxe -main HelloWorld -js HelloWorld.js
```

The compilation will result in the creation of the JavaScript source file *HelloWorld.js*, the contents of which are shown below:

```
(function (console) { "use strict";
var HelloWorld = function() { };
HelloWorld.main = function() {
    console.log("Welcome to Haxe Programming");
};
HelloWorld.main();
})(typeof console != "undefined" ? console : {log:function(){}});
```

The same Haxe code will be compiled into Python using the corresponding option in the Haxe compiler, as follows:

```
:  
haxe -main HelloWorld -python HelloWorld.py
```

This would generate the Python source code as output in the file *HelloWorld.py*, the contents of which are shown below:

```
class HelloWorld:
    @staticmethod
    def main():
        print("Welcome to Haxe Programming")
HelloWorld.main()
```

With the default installation of Haxe in the Windows environment, when you try to target certain programming languages like C#, there will be an error message like:

"Error: Library hxcs is not installed".

This can be rectified by using the following command:

```
haxelib install hxcs
```

After this installation, the source program will be compiled into the C# target with the following command:

```
haxe -main HelloWorld -cs HelloWorldCS
```

The point to be noted here is that 'HelloWorldCS' is a folder name, in contrast to the file names supplied for Python and JavaScript compilation. The above command creates a folder and places all the necessary files and sub-directories for the C# project. Another interesting feature is that the generated C# code is also compiled and an 'exe' in the name 'HelloWorld.exe' is built in the 'bin' folder, which will be readily executed.

The Haxe standard library

The standard library provided with Haxe has support for various commonly used tasks. It has three major categories as shown in Figure 3.

- The general-purpose component of the standard library covers features like arrays, strings, regular expressions and XML. It also covers advanced features like various encryption algorithms (*haxe.crypto*, *haxe.JSON*, etc).
- The system component has features like DB manipulation, file handling, process handling, etc.
- The target-specific library component covers features specific to the corresponding target languages.

Haxe based frameworks and tools

Haxe has good support for various domains as shown in Figure 4. There are various Haxe based tools which assist in the above-specified application domains:

- OpenFL
- Flambe
- Nape Physics Engine

The OpenFL framework is built using Haxe and

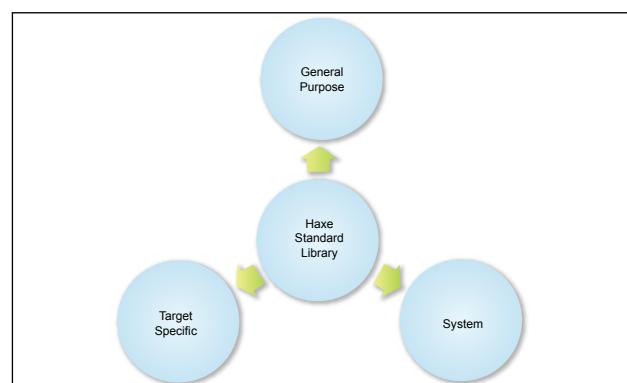


Figure 3: The categories in the Haxe standard library



Figure 4: Haxe applications

provides a Flash API for various mobile platforms. Flambe is a good gaming framework with support for mobile browsers. Nape is a physics engine fully written in Haxe. Similarly, there are various provisions for desktop application development, mobile application development and command line based applications.

HaxeUI

HaxeUI provides a good set of user interface elements, which can be used for building cross-platform interfaces. A sample (<http://haxeui.org/testapp.jsp>) UI using HaxeUI is shown in Figure 5.

Overall, Haxe provides a rich experience to developers in terms of its language support and wider array of target languages. The reliability of Haxe can be understood from the list of popular companies (Prezi, BBC, Toyota, etc) using it. This is listed in the official website of Haxe.

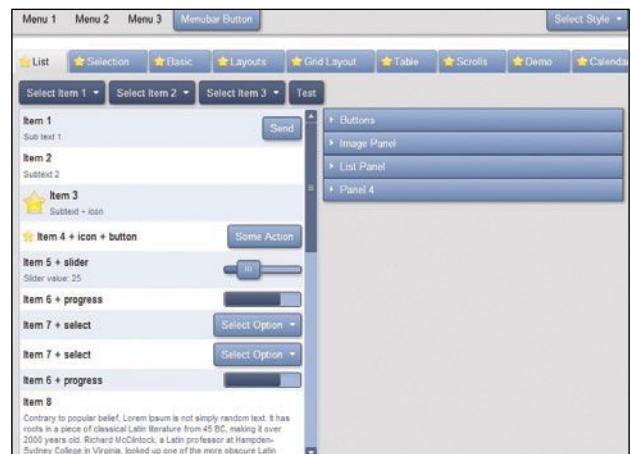


Figure 5: HaxeUI

To conclude, if you are a developer wanting to develop applications for a spectrum of platforms, you should definitely give Haxe a try. 

References

- [1] <http://haxe.org/>
- [2] <http://haxe.org/documentation/introduction/>
- [3] <https://github.com/HaxeFoundation/HaxeManual/raw/master/HaxeManual/HaxeManual.pdf>
- [4] <http://nekovm.org/>

Dr K.S. Kuppusamy

The author is assistant professor of computer science at the School of Engineering and Technology in Puducherry University. He has 10+ years of experience in academia and industry. His research interests include Web information retrieval, mobile computing and accessible computing. He has been awarded the Best Teacher Award at his institute, three times. He can be reached via mail at kskuppu@gmail.com.

OSFY Magazine Attractions During 2015-16

MONTH	THEME	BUYERS' GUIDE
March 2015	Open Source Firewall and Network security	SSD for Servers
April 2015	Web Development	Network Switches
May 2015	Virtualisation (containers)	Wireless Routers for SME
June 2015	Open source Databases	PaaS Solution
July 2015	Network Monitoring	MFD Printers for SMEs
August 2015	Mobile App Development	Hosting Solutions
September 2015	Backup and Data Storage	External HDD
October 2015	Programming Language	IaaS Solution
November 2015	Cloud Special	Firewall and UTMs
December 2015	Open Source on Windows	Online Backup solutions
January 2016	Android Special	Wifi Hotspot Devices
February 2016	Top 10 of Everything	External Storage

The Basics of MVC Architecture in PHP

MVC, which stands for Model-View-Controller, is a really good way to develop clean, scalable, powerful and fast applications in the least amount of time and with the least effort. Before exploring MVC, this article begins with a brief introduction to PHP.



PHP is an open source programming language that was developed in 1994 by Rasmus Lerdorf. PHP originally stood for *Personal Home Pages* but now stands for *Hypertext Preprocessor*, which is a recursive acronym. It is a loosely typed server-side scripting language used to build Web applications. A scripting language is a language that does not need to be compiled before use. And 'loosely typed' languages are those in which there is no need to define the type of a variable.

Today, more than 20 million domains use PHP, including

Facebook and Yahoo.

PHP is easily embedded with HTML, and is used to manage dynamic content and the databases of websites or, we can say, Web applications. We can use PHP with many popular databases like MySQL, PostgreSQL, Oracle, Sybase, Informix and Microsoft SQL Server.

How PHP works on servers

When a client (browser) sends a request for a PHP page to the server, the server reads the requested file from the disk

(storage) and sends this file to the interpreter. The interpreter then runs the PHP code, fetches the DB data (if required) and puts it into HTML tags. Once the interpreter completes its tasks, it sends the result back to the server, which sends this data to the client (browser) that made a request for this page.

MVC architecture with PHP

The Model-View-Controller concept involved in software development evolved in the late 1980s. It's a software architecture built on the idea that the logic of an application should be separated from its presentation. A system developed on the MVC architecture should allow a front-end developer and a back-end developer to work on the same system without interfering with each other.

Model

Model is the name given to the component that will communicate with the database to manipulate the data. It acts as a bridge between the View component and the Controller component in the overall architecture. It doesn't matter to the Model component what happens to the data when it is passed to the View or Controller components.

The code snippet for running *first_model.php* is:

```
<?php
class Model
{
public $string;
public function __construct()
{
    $this->string = "Let's start php with MVC";
}
}
?>
```

View

The View requests for data from the Model component and then its final output is determined. View interacts with the user, and then transfers the user's reaction to the Controller component to respond accordingly. An example of this is a link generated by the View component, when a user clicks and an action gets triggered in the Controller.

To run *first_view.php*, type:

```
<?php
class View
{
private $model;
private $controller;
public function __construct($controller,$model)
{
    $this->controller = $controller;
    $this->model = $model;
}
```

```
}
public function output()
{
    return "<p>" . $this->model->string . "</p>";
}
?>
```

Controller

The Controller's job is to handle data that the user inputs or submits through the forms, and then Model updates this accordingly in the database. The Controller is nothing without the user's interactions, which happen through the View component.

The code snippet for running *first_controller.php* is:

```
<?php
class Controller
{
private $model;
public function __construct($model)
{
    $this->model = $model;
}
}
?>
```

A simple way to understand how MVC works is given below.

- 1) A user interacts with View.
- 2) The Controller handles the user input, and sends the information to the model.
- 3) Then the Model receives the information and manipulates it (either saving it or updating it by communicating with the database).
- 4) The View checks the state of the Model and responds accordingly (lists updated information).

In the following code snippet, by running *first_example.php* we can see our MVC architecture work with PHP.

To run *first_example.php*, you must type:

```
<?php
$model = new Model();
$controller = new Controller($model);
$view = new View($controller, $model);
echo $view->output();
?>
```



By: Gaurav Kakran

The author is senior associate-Web development at Fierce Hound Media Pvt Ltd, Gurgaon. He is fond of developing Web applications in PHP, and can be reached at gauravkakran7@gmail.com

A Sneak Peek into PHP 7

PHP 7 is a new major version release, arriving 11 years after PHP 5. The scheduled release is before the end of the year. What will this new version have in its repertoire? This article gives the reader a taste of the things to come.



The good news for PHP developers is that PHP 7 RC5 (Release Candidate) is released on October 15, 2015.

PHP 7 is in the development phase, so do not use it in a production environment until the final version is out in November. To check when the next beta version will be released, visit <https://wiki.php.net/todo/php70#timetable>. The current stable version is PHP 5.6.14.

Let's now check out some of PHP 7's exciting new features.

The name and new Zend Engine

PHP developers have decided to choose version 7 instead of version 6 to avoid the confusion with past experimental projects. PHP 6 started as an experimental process earlier with a specific set of features. However, it did not reach the production phase. Hence, the developers decided to omit version 6 and instead name their new major release PHP 7.

The Zend Engine has been a core part of PHP since PHP 4. It's an open source execution engine that interprets the PHP language. With the PHP 5 series, the Zend Engine 2 was introduced with a built-in object model and performance enhancing capabilities. PHP 7 comes with the new engine version called PHPNG (Next Generation).

Improved performance

Though the current stable PHP 5.x series is working significantly faster than all its previous versions, the main goal of PHPNG is to get a much better and improved

performance compared to the PHP 7 with the new engine optimised memory usage and compiler does the compilation at prior to execution. Figures 1, 2 and 3 show the performance benchmarks provided by the Zend performance team.

[Reference: http://www zend com/en/resources/php7_infographic]

64-bit support

Though a part of the LAMP (Linux, Apache, MySQL, PHP) stack, PHP 7 introduces 64-bit support on Windows systems. It allows the running of both native 64-bit integers and large files on Windows systems.

New operators

PHP 7 comes with two new operators — the null coalesce operator and the combined comparison operator.

The null coalesce operator is denoted by double question marks (??). It returns the first operand if it exists and is not null; otherwise, it returns the second operand. It has the same usage as ternary operators in the C language.

```
<?php
$a = 10;
$b = $a ?? 20;
```

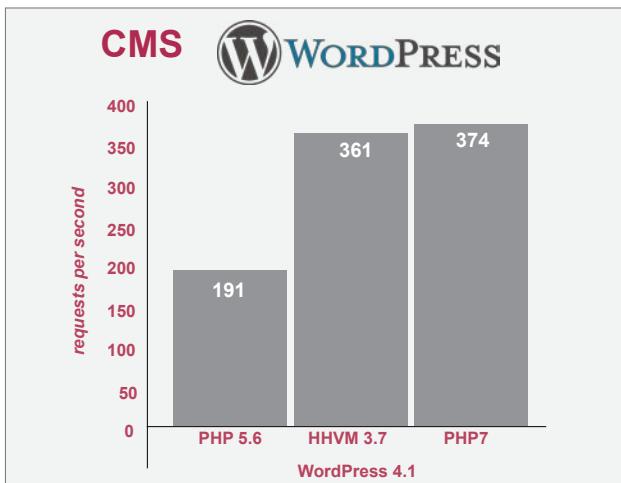


Figure 1: WordPress on PHP 7 (see Reference 2)

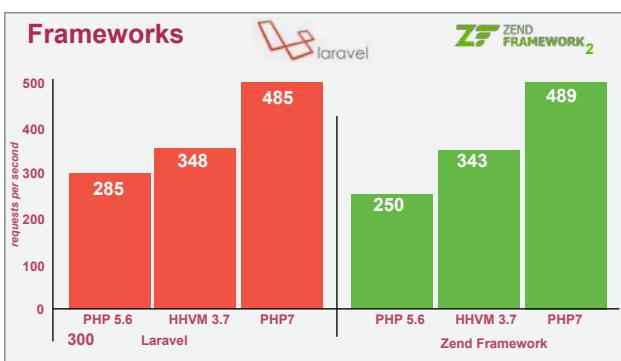


Figure 2: Frameworks performance under PHP 7 (see Reference 2)

```
//Equivalent to: $b = isset($a) ? 10 : 20
//isset() returns true if variable is set and not null.
```

The combined comparison operator, also known as the spaceship operator, is denoted by `<=>`. It returns zero if both the operands are equal, 1 if the left operand is greater and -1 if the right operand is greater.

```
<?php
$a <=> $b
```

The above expression will be resolved to 0 if \$a equals \$b, 1 if \$a is greater than \$b and -1 if \$b is greater.

Type declarations

Another feature introduced is to define the return type of functions and methods to avoid unintended return values. The code below states that `MyFunction()` returns a value which will be converted to bool automatically.

```
<?php
function MyFunction() : bool
{
    return 1;
```

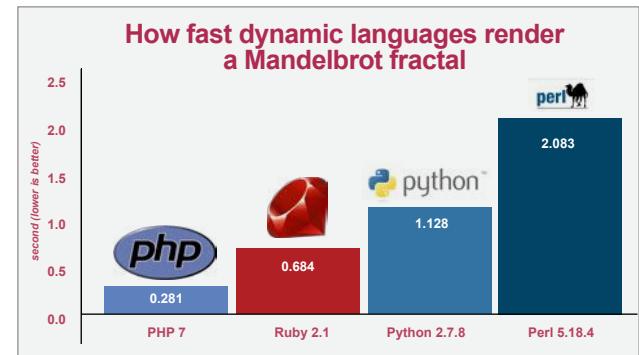


Figure 3: PHP 7 against other dynamic languages (see Reference 2)

```
}
```

```
echo MyFunction();
```

PHP 7 introduces four new types of declarations for scalar types — integer, float, string and Boolean. By default, in a scalar type, if we pass a float value to an integer parameter, it will re-correct it to the integer without outputting any warning or error.

```
<?php
function add(int $a, int $b) : int
{
    return $a + $b;
}
echo add(1, 2); // 3
echo add(1.5, 2.5); // 3 as floats are truncated by
default
```

Anonymous classes

Anonymous class (class without a name) is another feature provided which is already available in Java and other object-oriented languages. It has the same syntax as the normal class that was written in PHP but without a name. It can be used in case the class does not need to be documented.

```
<?php
var_dump(new class()
{
    public function Myfunction()
    {
        //code
    }
});
```

Import from the same namespace

This Group Use Declaration feature allows us to import many classes from the same namespace with a single line.

Continued on page 87...

Hybrid Drives: Get the Best of Both Worlds!



While hard disk drives were engaged in storing data on platters, the world moved on and solid-state drives made their appearance. SSDs, which are stacks of NAND flash memories, have their advantages as well as disadvantages. However, you can enjoy the best of both worlds by using hybrid drives, which incorporate both HDDs and SSDs.

Computer storage, specially the direct-attached type, began its journey around 1957 with IBM 325 systems, which were the first-of-a-kind ‘platter’ based disk drives available commercially. This led to people actually starting to consider storing data ‘persistently’ for their usage.

This revolution went through several milestones - advances in manufacturing processes and material sciences made the capacities of these ‘platter’ storage systems grow from megabytes to gigabytes in the 1990s, and to terabytes in 2007. Even today, these are the *de facto* first choice for personal computers, due to their potentially large capacity and high reliability.

A paradigm shift took place in this world of ‘rotating platters’ when the memory technology called ‘NAND flash’ emerged from Toshiba in 1984. In its initial form, NAND flash was intended to be used in handheld equipment for small capacity storage of microcode and data (EPROM like usage), but it soon became evident that we could stack up several such NAND flash chips together and still provide gigabytes of very fast drives in computers/laptops.

HDDs – Huge rotating storage

Taking a sneak peak into HDD design involves a deep dive into physics, kinetics and material sciences. Hard disk drives record the data by magnetising a thin film of ferromagnetic material. As depicted in Figures 1 and 2, a typical HDD design consists of a spindle that holds flat circular disks (or platters), which store the recorded data. The platters are made from a non-magnetic material (usually an aluminium alloy, glass or ceramic) and are coated with a thin layer of magnetic material (typically 10 to 20 nanometres thick). Data is recorded on these platters in concentric circles (called tracks, each containing several sectors of data) using a magnetic head or simply a head. Several such platters are put together and are made to spin at either 5,400 revolutions per minute or 7,200 revolutions per minute (and even higher in some HDDs) when the HDD is powered on. Data is binary encoded and read/written to the specific sector of the specific track of the specific platter using the head and actuator assembly driven by a motor.

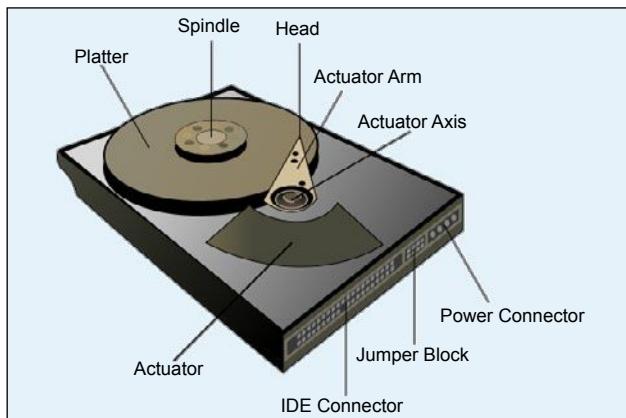


Figure 1: HDD1

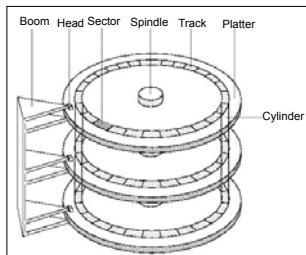


Figure 2: HDD2



Figure 3: Internals of a typical SSD

When the data at a specific address is accessed (read/written) by the host, the actuator moves the head such that the specific head is positioned above the required track of the required platter. Now the platter rotates, positioning the head above the specific sector of the track -- enabling the head to read/write the data from the required sector. A very interesting video on this is available at <https://upload.wikimedia.org/wikipedia/commons/f/f8/Harddrive-engineerguy.ogv>

Fun fact: Do you know that HDDs store the information using a quantum mechanical phenomenon called Giant Magnetoresistance? Read more about it at <http://www.cospa.ntu.edu.tw/aappsbulletin/data/18-6/18giant.pdf>. And yes, never bring a magnet near a rotating HDD. If you do, be prepared for a catastrophe!

SSDs – Fast and limited solid storage

SSDs or solid state drives consist of stacks of NAND flash memories (which essentially are plain semiconductor memories) put together. There are no rotating parts, which makes them light in weight, low in power consumption, noiseless in operation and fast in performance. But all these benefits do not come for free. NAND flash memory has an inherent disadvantage – it wears off as we keep reading or writing to it, and this wear is not very deterministic. Another characteristic of NAND flash is that unlike the conventional HDD, we cannot simply overwrite a data byte but we need to erase it first and then re-program it. Changing a byte from 0x55 to 0xAA involves changing it from 0x55 to 0xFF (that

is, erasing it) and then writing 0xAA. Moreover, the entire memory is divided into pages and blocks (typically 16K bytes => one page and 256 pages => one block), and we have to erase an entire block before we can program even a single byte in a single page of the block.

All this mandates very complex data block management and error correction algorithms to manage the data in the SSD drive firmware.

On top of all this, being semiconductor memory, NAND flash memory manufacturing costs are very high and, hence, the cost per GB of SSD memory is approximately 10-15 times more than that of a HDD. Thanks to research in material sciences and growing investments in NAND flash fabrication factories, the trends seem to show that SSDs will catch up with HDDs in the next three to four years.

Hybrid drives a.k.a. SSHDs – The best of both worlds

As we have just seen, both memory technologies have their own advantages and limitations. This has led to attempts to combine the best features of both. The engineering solution to this, pioneered by Seagate, is to use HDDs for bulk storage and have some NAND flash as cache, in front, making it an SSHD or solid state hard drive. From the host standpoint, an SSHD is a single intelligent disk drive which has two internal compartments – a NAND flash compartment like SSD and a magnetic storage compartment like HDD. This SSHD will intelligently pin down your most frequently used data (essentially, operating systems, programs and other such hot data) in the SSD (or NAND flash) cache, and will write your gigabytes of movies and multimedia to the HDD.

Figure 5 depicts a chart taken from the Seagate website, which compares the features of HDDs, SSDs and SSHDs.

How to make a choice

Choosing a drive for a laptop/PC would depend on the usage it would be subjected to and the budget. With all the learnings from the previous sections, we can definitely conclude the following.

An SSD is approximately 100 times faster than an HDD, making it a primary choice if you are looking for high speed accesses to all your data. It will enable your laptop to be very thin, light and would drain your battery to a much lesser extent. But then this would make the price higher and higher, as you seek more and more drive space. SSDs are typically suitable for business application laptops that need not store much (< 500 GB) but need to be extremely light, portable, and boot and perform very fast.

If you plan to store a moderate amount of data (around 1TB), but still need your boot and performance to be above average and are prepared to go slightly higher than the HDD price of the same capacity, SSHD is your best option. This would be true for typical high-end (home and business) laptops/desktops. How much flash to choose is again a decision

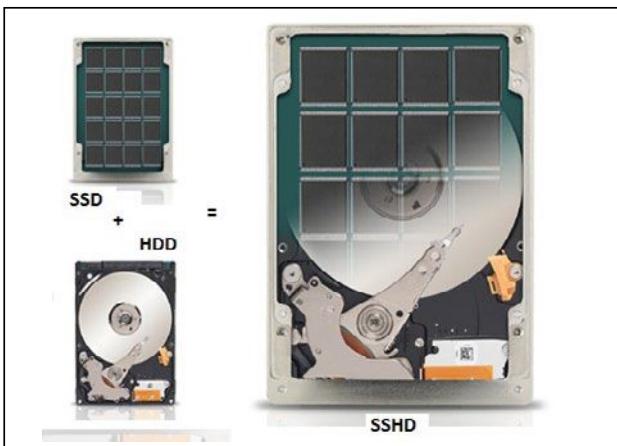


Figure 4: Concept of a typical SSD

Laptop Storage Selection Criteria						
HDD	★	***	***	***	**	**
SSHD	**	***	**	***	**	**
SSD	***	*	*	***	***	***

KEY: ★ - Good ** - Better *** - Best

Figure 5: A comparison of the features of HDDs, SSDs and SSHDs

Continued from page 84...

```
<?php
//current use syntax
use MyLibrary\Abc\classTest1
use MyLibrary\Abc\classTest2
use MyLibrary\Abc\classTest3

//group use declaration syntax
use MyLibrary\Abc{classTest1, classTest2, classTest3};
```

Engine exceptions

The ‘engine exceptions’ feature comes with PHP 7 to facilitate the error handling process in code. It allows us to replace fatal and recoverable fatal errors by exceptions. Catchable errors can be displayed and defined actions can be taken. In case the exception is not caught, PHP returns the fatal error. The `\EngineException` objects do not extend the `\Exception base class`. This results in two kinds of exceptions—traditional and engine exceptions.

```
<?php
function MyFunction($obj)
{
    //code
}
MyFunction(null); //Fatal error
```

Considering the previous code, the code below shows how a fatal error is replaced by an `EngineException`:

that has to be made carefully. For example, computer gaming enthusiasts should choose the one with more flash – others can do a budget tally and go for moderate flash configuration.

If you do not care too much about performance, power, form factor and the weight of the device, or are on a really tight budget or simply need loads of space (1TB to 16TB), you will have to stick to HDD. This is a good choice for a home/office desktop PC and low-end laptops.

In simple words, the choice is a standard one that every computer user and engineer has to make -- between capacity, cost and performance; and one wrong choice could drastically degrade your laptop/desktop experience.

References

- [1] <http://www.seagate.com/in/en/tech-insights/adaptive-memory-in-sshd-master-ti/>
- [2] <http://www.seagate.com/in/en/do-more/how-to-choose-between-hdd-storage-for-your-laptop-master-dm/>

By: Nilesh Govande

The author currently works at Seagate Development Centre, Pune. He has been working on different storage technologies for over 10 years and may be contacted at nileshgovande@gmail.com.

```
try
{
    MyFunction(null);
}
catch(EngineException $ex)
{
    echo "Exception Caught";
}
```

Lists of deprecated items have been removed from PHP 7 -- for instance, ASP style tags (`<%=`, `<%`, `%>`) and script tags (`<script language="php">`), safe mode and its `ini` directives, and so on. MySQL extension has been deprecated since PHP 5.5, and replaced by the mysqli extension (`mysqli_*` functions). Likewise, the `ereg` extension has been deprecated since PHP 5.3, and replaced with the PCRE extension (`preg_*` functions). Do visit https://wiki.php.net/rfc#php_70 for more information.

References

- [1] www.php.net
- [2] www.zend.com/n/resources/php7_infographic

By: Krunal Patel

The author is assistant professor in an engineering institute. His area of interest is open source technologies. He can be reached at krunalpatel27@yahoo.co.in.

Contribute to Shared Vehicles to Help Solve Urban Transportation Woes



In many cities, worldwide, commuters face many problems on a daily basis. International agencies like the UN are trying to get the world to cut down on pollution by reducing vehicular emissions. So, can the free and open source concept of technology come to the planet's rescue with a viable out-of-the-box solution to protect the environment and help the urban commuter?

The open source campaign has been a great and very successful technological movement. It's interesting to see the words 'technology' and 'movement' together, because they usually belong to completely different spheres of human activity. Because of the efforts of pioneers like Richard Stallman and Linus Torvalds who worked against the odds and firmly established legal, technological and, most importantly, the cultural foundations of the open source movement, we are fortunate to enjoy the fruits of the hard work of thousands of hackers, free of cost. The Linux operating system is a fine testament to what can be achieved by volunteers working in a spirit of cooperation and sharing knowledge freely. Free and open source software (FOSS) is now firmly established in the sense that it is no longer battling for survival, but for improvement, and there are loyal communities of hackers and users around the world who keep the FOSS world running.

Practising openness beyond software

If we look closely, this is not just a technological

phenomenon, but something much more than that. It is very much like a social movement involving the fight for freedom, justice and equality. It is also a cultural movement that practices and encourages questioning, creativity and volunteer participation. Because of this tremendous success, the FOSS movement started having a significant impact on other spheres. The first and obvious impact was on software businesses, with many companies changing their business models in which *production* was no longer in the hands of just a few in the company. Because electronics and computer science are very closely related, there has also been an open hardware movement, with many successful projects like Arduino. Other initiatives include making some government data public under the *open data* policy by many governments around the world, so that researchers, social activists, the public and software developers can use the data in their own way. We even have an *Open Source Drug Discovery* project. So where else can we apply this idea of 'open source'? Can we solve the urban transportation problem, for example?

The urban transportation problem

The urban transportation problem is something that almost every one of us experiences. Particularly in cities like Bengaluru, just visiting three or four places in different parts of the city can consume a whole day. Apart from the waste of time for an individual, imagine the amount of productive time lost for so many in the city. It's immaterial whether you own a private vehicle or depend on the public transportation – productive working hours are unnecessarily lost.

Many government and private players are trying to solve the problem in their own way, and often, these efforts are not coordinated. For example, we may have a metro but many people won't use it for the simple reason that it doesn't provide an end-to-end solution. That is, if it takes more than a few minutes to walk to the metro station from your house, you may have to depend on other transportation methods and this complicates the problem. Though city buses are very convenient in terms of an 'end-to-end' solution, they consume too much time.

And then we have the problem of air pollution. Electric vehicles can solve this problem, but we have two critical problems—the limited range of the vehicle and the time needed for a commute.

In such a context, can we apply or use ideas from the FOSS movement to solve urban transportation problems?

What would be the key activities? What contribution can we expect from the volunteers? How can the open source movement actually help in such a traditional industry involving automobiles and transportation?

Project Vidyut: A platform for electric vehicles

Project Vidyut is a not-for-profit initiative that aims to solve the transportation problem. It is greatly inspired by the ideals of the FOSS movement. Its focus is on electric vehicles. And there are two specific approaches:

1. Designing and building electric vehicles and improving key parameters like range and charge duration.
2. Designing 'intelligent' systems whereby we can use existing electric vehicle technology.

In the first approach, we'll have to design electric vehicles (whether two-wheeled or four-wheeled), do all the engineering analysis, and coordinate the efforts of different research/engineering teams to solve critical battery problems.

In the second approach, we will not try to improve the design of vehicles, but attempt to design a system that uses vehicles with existing technology. Basically, the idea is to build an 'intelligent' system that uses a combination of different modes of transportation, according to the situation, availability and cost. This will not be another system designed in isolation, but more like *stitching* things together in real-time.

FEEDBACK FEEDBACK

We Value Your Feedback

We love to hear from you as **OSFY** consistently strives to make its content informative, interesting and engaging.

Please share your feedback/ thoughts/ views via email at osfyedit@efy.in



We welcome your comments/ suggestions and encourage you to send them to:
The Editor, D-87/1, Okhla Industrial Area, Phase 1, New Delhi-110020.

FEEDBACK FEEDBACK

The SharedVehicles Project

The idea behind SharedVehicles is to build a system that's very dynamic, can respond to changes in demand and is cost effective for everyone; it has to be a shared system. Of course, owned vehicles will still be a part of the system, but they can't take you far if we've to make the system fast and responsive beyond a point. Consider the following hypothetical situation.

Let's suppose you've got to go from A to D, with the route being A-B-C-D. Imagine that there is a mass transportation system or metro available from B-to-C. If you're at A, to go to B, you can take a 'shared' electric bike, which will be available at A. Then you drop the bike at B, take the metro to C, and you take another electric bike parked at C, to reach D, and leave it there for someone else to pick up.

Of course, in reality, things will be much more complicated than this, and here, I'm simply suggesting one possibility.

If we've to implement such a system, one problem we would face is understanding the requirements of users and the corresponding behaviour of the system. As the system becomes too complex, we may have to employ some powerful concepts from the world of simulation. In such a situation, we can employ agent based modelling and simulation. So what exactly is agent based modelling? How is it useful for urban transportation? And how is it related to the open source movement?

Agent based modelling

All of us have attended physics classes in high school (and for many, in college, too), where the physics teacher started the discussion saying, "Consider a body of mass...". For a physicist, the object of study was a *body*. Similarly, when an artificial intelligence researcher is trying to simulate the behaviour of a person or an intelligent being, he will call either of them an agent. But note that the agent need not be as intelligent as a human—it can be very elementary in terms of behaviour or level of intelligence. Usually, agents have a set of rules governing them, and they behave according to those rules to achieve some objective. Typically, we study the behaviour of an agent in relation to its surroundings or an environment. Sometimes, the agent will be alone, while most of the time agents interact with other agents, and such a setting is called *multi-agent simulation*.

There are many frameworks available to implement such agent based modelling. Many of them are written in Java. For example: Mason - <https://cs.gmu.edu/~eclab/projects/mason/>, Jade: <http://jade.tilab.com/>, etc.

Mason is easy to get started with, to understand, and of course, to manipulate. It is a 'single-process, discrete event simulation library' according to official documentation. It also comes with a visualisation library. Simulation can be run without visualisation, but the latter often helps us to understand simulation better, and can be intelligently adopted to a particular simulation.

Examples of simulation provided in the Mason framework demonstrate the capability of agent based modelling. In one

simulation called *Ant foraging*, the behaviour of ants is modelled. Each ant is an agent that acts on its own; however, it is governed by some rules. Ants discover food, and establish a food trail between the source and the nest. In another simulation called *Virus Infection* we can see the spread of a virus. There are infected individuals or agents, and uninfected individuals or agents, and the simulation studies how the virus spreads. There is also the famous example – John Conway's 'Game of Life'.

It's easy to run Mason. Just run the command `java -jar mason.19.jar` once you've downloaded Mason from the site mentioned above. Mason has very good documentation in terms of a manual - <https://cs.gmu.edu/~eclab/projects/mason/manual.pdf>. The best way to learn Mason is through the examples and tutorials that are provided.

Simulation and the SharedVehicles Project

Now that we've understood the concept of agent based modelling and its possible applications, we can look at how to use it in the SharedVehicles Project. Here passengers, vehicles, buses, the metro, taxis and private vehicles are all *movable*. There can be a *movable* interface or a class. Then we can implement/extend it to give behaviours specific to buses, passengers, metros, etc. We can then go one crucial step further by adding parameters like costs, charge for the electric battery, locations of charging stations, etc, and try to obtain the cost for each person and the profit/loss for different operators. When we reach a point at which we can show that for a specific arrangement, the cost of transportation is much less than what passengers are paying today, and where profits for the operators can be substantial, then we would have achieved a breakthrough. That would create an incentive for the government/entrepreneurs to implement such a system.

To put it in a different way – anybody can contribute ideas in terms of code. Just like any other FOSS project, this will be available to everyone. When the simulation suggests that there can be a better, faster and more efficient system that can be implemented, which can also become profitable, that system will get implemented by a willing entrepreneur.

The SharedVehicles project is available at <https://github.com/project-vidyut/>, where you can also find links to the resources needed to implement the project. You can visit Project Vidyut site: <http://www.projectvidyut.org/> This is a new experiment in which we're trying to solve a burning real-world problem with ideas and experience from the FOSS movement. I request you to contribute in terms of code, ideas, and constructive criticism. 

By: Mahesha Hiremath

The author is an independent data analytics consultant from Bengaluru. He is building a platform for electric vehicles called Project Vidyut, a not-for-profit initiative for promoting the use of electric vehicles, and also to build open source electric vehicles. He can be reached at mahesha-at-projectvidyut-dot-org and tweets with [@MaheshaHiremath](https://twitter.com/MaheshaHiremath).

What You Should Know about Open Source Licensing

FOSS is intellectual property and therefore is an asset that should be protected. The owner of the FOSS has a right to decide and control how it can be used by other developers and the public, in general. Various types of FOSS licences are used by software developers.



Many FOSS developers and designers want to release their work as open source projects. They also want to get recognised for their work. The open source community is vibrant because the essence of it is preserved through multiple iterations of the code. But there is a lot of confusion out there about what licensing means exactly. Licences fall in the domain of the law, which means matters related to it quickly become relatively complex. This article can't be considered hard core legal advice, but you may find it useful when talking with your lawyer or, if necessary, coming to a decision without the help of a lawyer.

Differences between free software and open source software

To understand open source licensing, we need to understand the difference between free software and open source software. The term, FOSS or Free and Open Source Software shows that both are different entities. Richard Stallman and others started the Free Software Movement. The definition

of free software is focused on the freedom to share code with your neighbour. It deals with the moral or ethical part of source code sharing. The point to be noted here is that not all the licences approved by open source proponents meet with the approval of free software advocates.

The open source philosophy is based on pragmatism rather than ideology. It promotes the notion that "...advocacy targeted at businesses should emphasise the technical merits and profitability of open (free) development models, rather than talking about ethical or political issues" (The Open Source Initiative; www.opensource.org). Followers of this movement prefer to base their arguments on the economic and technical merits of making source code freely available, rather than the moral and ethical principles that drive the Free Software Movement.

So, coming back to licences again, the Free Software Foundation uses a specific licence under which it provides software. The Open Source Initiative (OSI) seeks support for all open source licences, including the one from the Free



Figure 1: Open source versus free software

Software Foundation. The OSI is an organisation that focuses entirely on gaining support for open source software. The idea behind the OSI is to get more companies behind open source by allowing them to write their own open source licences, and have those licences certified by the Open Source Initiative (<http://opensource.org/licenses/index.html>).

GNU licences

The Free Software Movement is headed by the Free Software Foundation, a fund-raising organisation for the GNU project. The GNU General Public License (GPL) is a means of implementing a concept called copyleft, which attempts to negate copyright for the purposes of collaborative software development.

GPL (V2 or V3) is a copyleft licence that requires anyone who distributes your code or a derivative work to make the source available under the same terms. V3 is similar to V2, but further restricts use in hardware that forbids software alterations.

The GPL, in brief, provides that anyone distributing a compiled program which is licensed under the GPL must also provide source code, and is free to make modifications to the program as long as those modifications are also made available in source code form. This guarantees that once a program is ‘opened’ to the community, it cannot be ‘closed’ except with the consent of every author of every piece of code (even the modifications) within it.

Most Linux programs, Git, WordPress, etc, are licensed under the GPL.

There are two versions of GPL in use— GPLv2 and GPLv3. The latter contains compatibility regulations that make it easier than before to combine GPL code with code that was published under different licences. GPLv3 also addresses digital rights management. It addresses what would happen if the terms of the licence are violated and specifies the cure for the violation.

To use the GPL, you have to include certain information in the software’s code, along with a copy of the licence. Further details and the licence itself can be found at www.gnu.org/licenses/licenses.en.html.

Another flavour of GPL is the GNU Lesser General Public License (LGPL). Because GPL requires other software with parts of licensed code to also be licensed under the GPL, developers cannot use GPL-licensed code for paid and proprietary software. The LGPL negates this aspect of the GPL since it does not require other projects with parts of the code to be similarly licensed. It grants fewer rights to a work than the standard GPL.

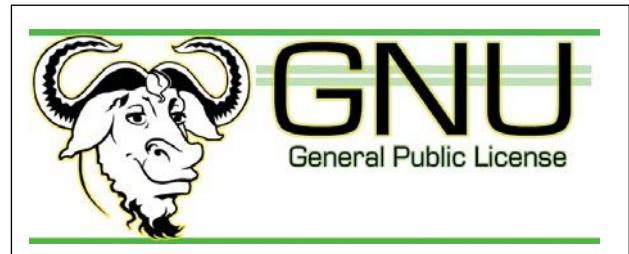


Figure 2: GNU–GPL logo

BSD and the Mozilla License

Other open source licences like the BSD and Mozilla License also allow redistribution, but have fewer rules regarding the conditions for redistribution. The BSD License allows anyone to redistribute the work or any derivative without any source, if such is the desired path. Some people believe that these licences grant users more freedom, since they include a ‘superset’ of the rights granted by the GPL. Other people believe that only the GPL truly respects users’ freedoms by guaranteeing that no one can co-opt their original work.

A ‘permissive’ licence is simply a non-copyleft open source licence — one that guarantees the freedoms to use, modify and redistribute, but permits proprietary derivative works. BSD is a permissive licence that lets people do anything with your code with proper attribution and without warranty.

The Mozilla Public License (MPL 2.0) is maintained by the Mozilla Foundation. This licence attempts to be a compromise between the permissive BSD License and the corresponding GPL.

Apache License

The Apache License is also permissive, and allows express grant of patent rights from the contributor to the recipient. Because some licences can be applied only to copyrights and not patents, this flexibility would be an obvious factor in a developer’s choice of licence. Redistributing code also has special requirements, mostly pertaining to giving proper credit to those who have worked on the code and to maintaining the same licence. Also, the Apache License states that no one can take these rights away once they’re granted. In other words, you don’t need to worry that after

you’ve created some awesome derivative of the licensed code, someone down the road will swoop in and say, “Sorry, you can’t use this code any more.”



Figure 3: BSD License



Figure 4: Mozilla License

MIT License

The MIT License is a permissive one, the wording for which is short and to the point. It lets people do anything they want with your code as long as they provide attribution back to you and don’t hold you



Figure 5: The Apache Software License



Figure 6: MIT License

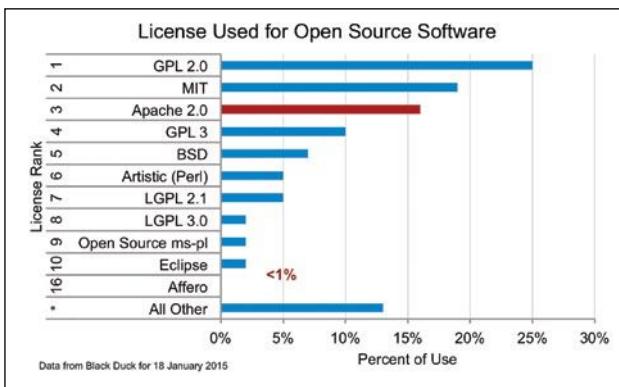


Figure 7: Licences and usage statistics

liable. The MIT License is the least restrictive one. It basically says that anyone can do whatever they want with the licensed material, as long as it's accompanied by the licence. You can use, copy and modify the software however you wish to. No one can prevent you from using it on any project, from copying it however many times you want and in whatever format you like, or from changing it however you want to.

There are also shared source licences that have some similarities with open source, such as the Microsoft Reciprocal License (MS-RL). These are mainly used by Microsoft and can range from extremely restrictive to being comparable with the GPL.

Not choosing a licence for open source software

Opting out of open source licences doesn't mean you're opting out of copyright law. The absence of a licence means that default copyright laws apply. This means that you retain all rights to your source code and that nobody else may reproduce, distribute, or create derivative works from your work. You retain all rights and do not permit distribution, reproduction or derivative works. You may grant some rights in cases where you publish your source code to a site that requires accepting terms of service. For example, publishing code in a public repository like GitHub requires that you allow others to view and fork your code.

For more information, do refer to the OSI website at <http://opensource.org>. 

By: Sricharan Chiruvolu

The author is a FOSS enthusiast and an aspiring computer science engineer. He's currently studying for his computer science degree in Bengaluru. He loves to explore new open source technologies. You can reach him via email at mail@sricharan.xyz.

An
EFY GROUP
EVENT

INDIA ELECTRONICS WEEK

January 11-13, 2016 • Bengaluru

I Loved
'Electronics Rocks 2014'

So the 2016 edition
is a must for electronics
professionals like me.

Ramaswamy Iyer,
Sr Design Engineer

4th Edition

electronics Rocks 2016

January 11-13, 2016 • BIEC, Bengaluru

2000-plus electronics design professionals and enthusiasts visited Electronics Rocks 2014.

The next edition (in 2016) will be held as part of India Electronics Week, so the event will be even more action-packed.

For more details, visit: www.electronicsrocks.com

A Brief on Search, Frames and Windows in GNU Emacs



We have been carrying articles on GNU Emacs for quite some time now. This is one more article in the series.

In this next article in the GNU Emacs series, we shall learn how to perform text search in a buffer, and introduce the concept of windows and frames.

Search

You can copy the following poem, which I wrote in 2013, in the *scratch* buffer or a file inside GNU Emacs to try out the search commands:

Emacs is an operating system
Which unlike many others, is truly a gem
Its goodies can be installed using RPM
Or you can use ELPA, which has already packaged them

You can customise it to your needs
You can also check EmacsWiki, for more leads
Your changes work, as long as reload succeeds
And helps you with your daily deeds

People say it lacks a decent editor
But after using its features, they might want to differ
Using Magit's shortcuts, you might infer
That it is something, you definitely prefer

Plan your life with org-mode
You don't necessarily need to write code
TODO lists and agenda views can easily be showed
Reading the documentation can help you come aboard

Emacs is a double-edged sword
Its powerful features can never be ignored

Customisation is possible, because of free software code And this is my simple ode.

You can search for a word in a buffer using the *C-s* short cut. You will then be prompted with *I-Search*: in the mini buffer, where you can type any text and GNU Emacs will try to find words matching it, in the buffer. This is an incremental forward search and is case insensitive. Thus, if you search for the word ‘todo’ in the poem, it will match the string ‘TODO’. You can exit from the incremental search using the *Enter* key, or abort the search using the *C-g* key combination. If you want to do an incremental search in the reverse direction - from the cursor position to the top of the buffer -- you can use the *C-r* short cut.

If you place the cursor on the letter ‘E’ in ‘Emacs’ in the poem’s first line, and press *C-s C-w*, Emacs will try to find all occurrences of the word ‘Emacs’ in the text. Suppose you have cut or copied text to the kill ring, you can search for this text by using the *C-s C-y* short cut. You can repeat the previous forward search using *C-s C-s*, and the previous backward search using *C-r C-r* short cuts.

The first occurrence of a text can be looked up in the forward direction using *C-s*. This will prompt you in the mini buffer with a *Search*: string, where you can type the text that you want to search for, and then press the *Enter* key. It will then search for the text and move the cursor to the matching word. This is a non-incremental forward search. Similarly, you can perform a non-incremental backward search using *C-r*. You can then input the search string to be searched for, followed by the *Enter* key.

Regular expression searches are very useful too. In order

to search forward for an expression, you can use *C-M-s* followed by the *Enter* key, which will prompt you in the mini buffer with the string ‘Regexp search:’. You can then enter a regular expression. This will only match the first occurrence of the text and the search will then terminate. You can perform a one-time backward regular expression search using the *C-M-r* short cut. To perform an incremental forward search, you need to use *C-M-s* and you will be prompted with the string ‘Regexp I-search:’, where you can provide the pattern to match. For example, ‘[a-z]+-[a-z+]’ will match both the expressions ‘org-mode’ and ‘double-edged’ in the poem. You can use *C-M-r* for an incremental backward regex search.

A common use case is to find and replace text in a buffer. The sequence to be used is *M-x query-replace* followed by the *Enter* key. You will then be prompted with the string ‘Query replace:’ where you will be asked which word or phrase is to be replaced. For example, if you mention ‘ode’, it will again prompt you with ‘Query replace ode with:’ and then you can enter the replacement string. You can also search and replace text by matching a regular expression with the *C-M-%* shortcut key combination.

Frames

The outermost user interface boundary of GNU Emacs is called a frame. In fact, when you split the GNU Emacs user interface, you are actually creating windows. So, in GNU Emacs, you have windows inside a frame. This is in contrast to today’s user applications, where the entire application is contained in a ‘window’. This is important terminology to remember when using GNU Emacs.

You can create a new frame using the *C-x 5 2* key combination. You can move the cursor to the next frame using *C-x 5 o* (the letter ‘o’), and delete the current frame using the *C-x 5 0* (zero) shortcut. This will not delete the existing buffers, but only the view. In order to open a file in a new frame, you can use *C-x 5 f*. You can also open a file in a new frame in read-only mode using *C-x 5 r*. To switch to the buffer in a new frame, use the *C-x 5 b* key combination.

Windows

You can split a frame vertically to create two windows using *C-x 2* (Figure 1).

To split horizontally, you can use *C-x 3* (Figure 2).

To move the cursor to the next window, use *C-x o* (the letter ‘o’). You can delete the current window using *C-x 0* (zero). Note that this does not delete the buffer, but just the view. If you have multiple windows, want to retain the current window and remove the rest of the windows from the display, you can use *C-x 1*.

You can open a file in a new window using *C-x 4 f*. You can also select an existing buffer in another window using *C-x 4 b*. If you have multiple windows that you would like to

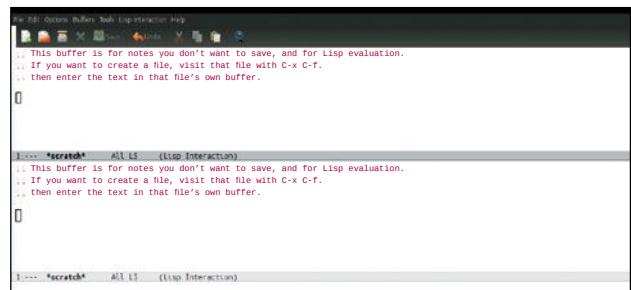


Figure 1: Split frame vertically

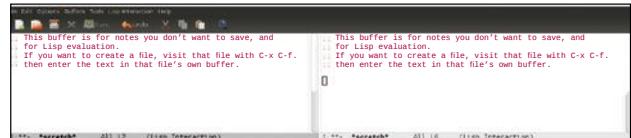


Figure 2: Split frame horizontally

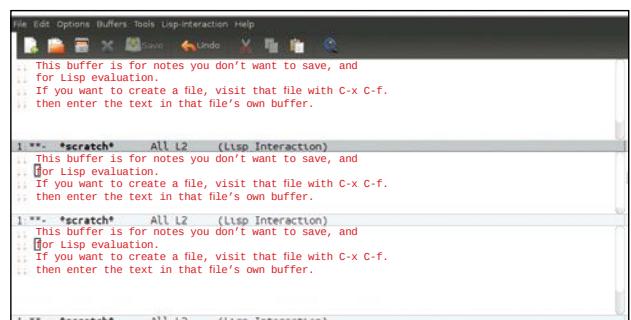


Figure 3: Balanced windows

be balanced equally, you can use *C-x +*. Figure 3 shows an Emacs screenshot with three windows that are balanced.

You can scroll the contents in the other window using *C-M-v*. You can scroll backwards using *C-M-Shift-v*.

You can also use the following shortcuts in your *~/.emacs* to simplify the shortcuts used to split and remove windows:

```
(global-set-key (kbd "C-1") 'delete-other-windows)
(global-set-key (kbd "C-2") 'split-window-below)
(global-set-key (kbd "C-3") 'split-window-right)
(global-set-key (kbd "C-0") 'delete-window)
```

If you would like to make a window wider, you can use the *C-x }* shortcut and, to reduce it horizontally, you will need to use *C-x {*. You can use a prefix count to perform the operation ‘n’ times. For example, use *C-u 5 C-x {* to shrink a window horizontally. To make a window taller, you can use the *C-x ^* short cut; and to make it smaller, you have to use a negative prefix, for example, *C-u -1 C-x ^*.

By: Shakthi Kannan

The author is a Free Software enthusiast and blogs at shakthimaan.com.

Speech Signal Analysis Using Praat

The word 'praat' means 'talk' or 'speak' in Dutch. This article introduces the reader to Praat, a rich, feature-laden piece of software that helps to analyse speech signals. It covers phonetics and various types of speech analyses.

Speech plays the most important role during verbal communication between humans. And phonetics is the study and classification of speech sounds. It is basically categorised into the following three groups:

1. Articulatory phonetics deals with speech production and how sound is produced.
2. Auditory phonetics deals with speech perception, and how sound is perceived.
3. Acoustic phonetics deals with the study of properties of the speech wave.

This process is termed the speech chain or the speech communication pathway.

An introduction to speech processing

Analysis of speech signals is an important task. It deals with the manipulation of sound signals to extract meaningful information from them. There are many techniques for speech processing and, in this article, we will discuss some basic methods. Interested readers can go through an excellent tutorial on Praat at <http://www.fon.hum.uva.nl/david/sspbook/sspbook.pdf> (Reference 1).

Speech, being a non-stationary signal, continuously keeps on changing; hence, in order to model the speech signal, we follow the strategy of segmentation, which is the process of assuming the speech wave to be a static signal for a short period of time in which it remains almost constant. The typical length of such intervals is 20ms to 30ms. Hence, a speech signal is segmented into frames of 20ms-30ms, and each frame is analysed separately.

Spectrum analysis

Normally, the signal is represented in the time domain, where the horizontal axis represents time and the vertical axis represents the amplitude (loudness) of the speech signal (i.e., amplitude → time). You can look at the upper half of Figure 3, which is a time domain signal. The difficulty with the time domain is that it can provide only a limited amount of information about the signal. Fourier analysis is an important process, which transforms the signal from the time domain to the frequency domain, also referred to as the Fourier domain. The frequency domain represents the signal in amplitude (decibel) → frequency (Hertz) dimensions. The signal representation in the frequency domain is also known as the frequency spectrum, or just the spectrum of the signal.

	1209 Hz	1336 Hz	1477 Hz	1633 Hz
697 Hz	1	2	3	A
770 Hz	4	5	6	B
852 Hz	7	8	9	C
941 Hz	*	0	#	D

Table 1: DTMF standards

A simple time domain signal (with a single frequency f) shows a single spike at f Hz in the frequency domain. A composite signal (combination of multiple frequencies) presents quite a complex spectrum. Let us look at the following simple example—Dual Tone Multi Frequency

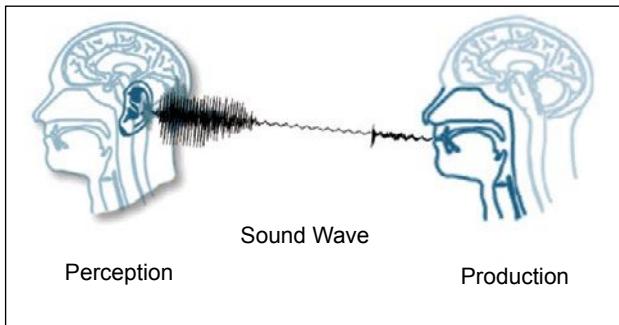


Figure 1: Speech chain (Source: The Virtual Linguistic Campus (VLC) course on Speech Science)

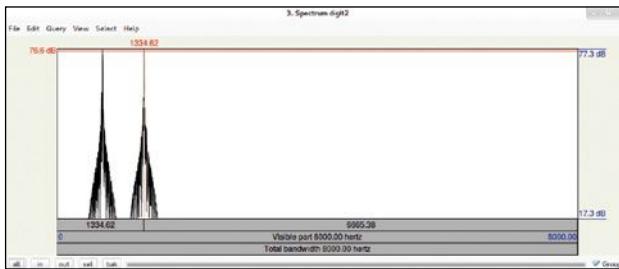


Figure 2: Spectrum for digit '2'

(DTMF) is the pair of frequencies assigned to each digit when we dial some number through the phone. Table 1 lists such frequencies corresponding to each digit. This is helpful in determining the dialled number. By Fourier analysis, we can obtain the frequencies present in a specific sound signal (dial tone). Figure 2 gives the frequency spectrum for the tone generated by pressing a specific digit.

In addition to DTMF, spectrum analysis is used to identify musical instruments, vowel identification and gender recognition, from audio signals.

Spectrogram analysis

A spectrogram represents three dimensions of a signal, i.e., time, frequency and amplitude. The horizontal and vertical dimensions indicate time (sec) and frequency (Hz) values, respectively. The darkness of the bands indicates the energy density or strength of the amplitude. Absence of darkness (pure white colour) represents the silence zone.

Spectrogram analysis is widely used in vowel identification, silence detection or formant analysis from specific speech utterances. Further, a spectrogram can also be used to identify the category or class of sounds (such as nasals, plosives, fricatives, etc). A typical spectrogram is shown in Figure 3.

Formant analysis

A formant can be considered as a resonance of the vocal tract. In a typical spectrum envelope of a speech signal, formants depict picks of the envelope. Most of the vowels are characterised by the first two formants (F1 and F2) only.

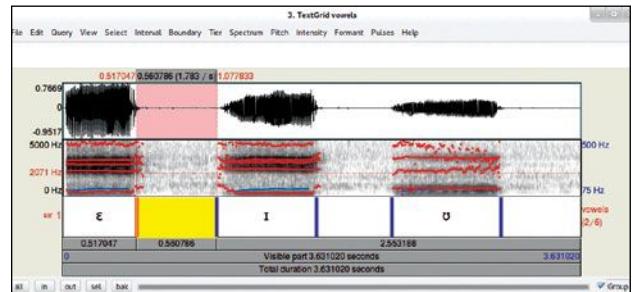


Figure 3: Spectrogram for three vowels

This is because formant frequencies reflect the articulatory gestures related to vowel height and the front-back position.

Format analysis is widely used in linguistics. There are many techniques to extract formants from a given speech signal. Let's discuss the easiest approach of getting Formant details from spectrograms. View the spectrogram of the signal and click on *Formant* >> *Show formants*, which will show the formant transitions in the spectrogram in red colour, as shown in Figure 3.

Praat

Praat is a tool designed for speech analysis. It was developed at the University of Amsterdam by Paul Boersma and David Weenink. According to them, Praat is a tool for “doing phonetics by computers.”

Downloading and installing Praat

There are multiple ways of getting Praat. Let us discuss the methods of downloading it on an Ubuntu 13.10 machine. However, the process remains almost the same for all Linux distributions.

1. Standard way (through terminal):

```
$ sudo apt-get install praat.
```

2. Get it from the Ubuntu Software Center.
3. Get it from the Synaptic Package Manager.
4. Download it from the Praat website (Reference 2) and extract the archive.

Figure 4 shows a screenshot of Praat installed on the machine. Double clicking the Praat icon will open it. Alternatively, on a terminal, just typing *praat* will open its interface. The Praat icon comprises lips and ears, which exhibit the main speech production and perception organs in the human body.



Figure 4: Praat after installation

Getting familiar with the Praat environment

Once launched, two windows are opened in Praat. One is the *Object* window and the other is the *Picture* window. We will discuss only the *Object* window, since most of the basic functionalities can be achieved through this interface alone. Praat is special in that it offers dynamic buttons upon selection of different sound objects. We will quickly go through the important buttons and their features. Readers are encouraged to explore the features of the rest of the buttons themselves.

Menu	Functionality
Praat	New and open Praat script (deals with creation and opening of Praat scripts)
New	Records sounds with different options and creates a text grid for annotations
Open	Deals with reading different sound files and loading them into <i>Object</i> window
Save	Deals with saving the recorded and processed sound objects into the desired file types

Table 2: Praat main menu showing key functionalities

It is very easy to record a sound using the *New >> Record Mono Sound* menu option. Selecting this option presents a window that is ready to record the sound with some default set parameters. You can modify sampling frequency or channel types. If you do not understand the terms, then keep the default settings unchanged. Pressing the *Record* button starts recording until the *Stop* button is pressed. You can test your recording by clicking the *Play* button. If the recording is acceptable, then specify some meaningful name to the recorded sound file and click on *Save to list* and then on *Close*. The sound object will now appear in the *Object* list. If the recording is not satisfactory, then repeat the recording. A typical screenshot of the Praat Sound Recorder is shown in Figure 5. Note that recording the sound and saving it saves the sound file as the object in the current session only. If you want to permanently save the recorded sound, then use *Save >> Save as WAV file...* menu.

Various functionalities

- o *Reverse a sound*: To record or read a sound in the *Objects* window, select the sound object which you want to reverse. Click on *Modify >> Reverse*, and play the sound once again. It will play the sound in reverse. To save the reverse version of the sound, use the *Save* menu from the top panel. Performing the same operation again will get you back to the original sound.

- o *Text-to-speech synthesis*: You can create a speech synthesiser from a given text. Click on *New >> Sound >> Create SpeechSynthesizer*. Select the language and voice variant (male/female) and click on the *Play Text* button.

- o *Amplify a sound*: A soft sound can be made louder by scaling its amplitude. To do this, select the target sound object, and click on *Modify >> Formula*. In the window that

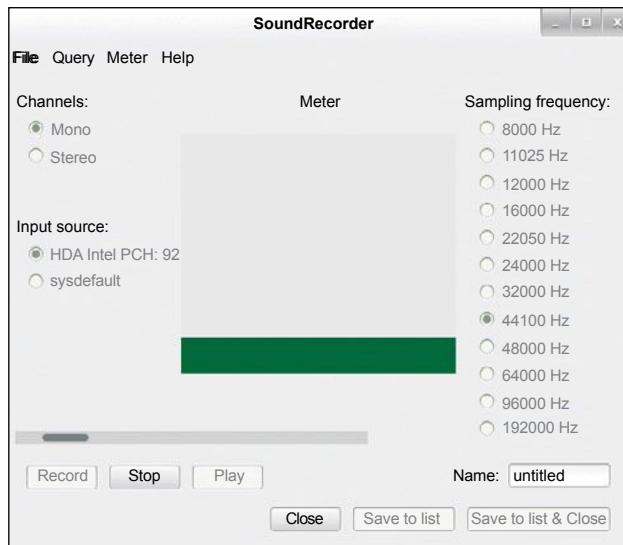


Figure 5: The Praat sound recorder

appears, enter the scaling factor for the sound. For example, *self*2* will double the amplitude of the sound.

Scripting in Praat

A script comprises a set of commands that need to be executed in a specific sequence to perform a certain task. Praat also supports scripting. Users can write scripts, which are interpreted by an interpreter.

Executing scripts

Let us start with a simple script. Click on *Praat >> New Praat Script* to open the script editor. The following are some of the usual commands that are often needed in scripts.

Command	Task
<i>Play</i>	Plays a selected sound object
<i>Remove</i>	Removes a selected sound object from the <i>Object</i> window
<i>writelnInfoLine:</i>	Prints the message passed as the argument
<i>appendInfoLine:</i>	Appends the message into the existing display rather than overwriting
<i>Clearinfo</i>	Clears the Praat <i>Info</i> window, where the output is displayed
<i>runScript:</i>	Executes an external script

Table 3: Some common scripting commands

Dealing with forms

Praat facilitates the use of forms to get input from users at runtime. Let us look at the use of forms with an illustrative example.

Example 1: To get personal details from the student (such as name, age, gender and results) through the form, and display these details, type:



Figure 6 (a): Form display after script execution

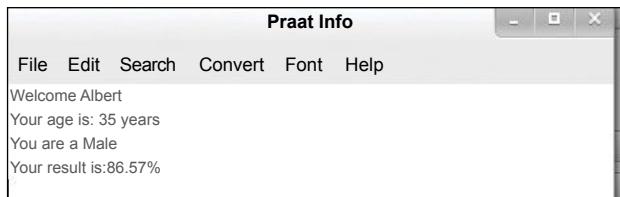


Figure 6 (b): Output

```
form Submission Form
    comment Please enter following information
correctly:
    sentence name:
    integer age:
    choice gender: 1
        button Male
        button Female
    real result:
endform

writeInfoLine: "Welcome ",name$
appendInfoLine: "Your age is: ",age," years"
appendInfoLine: "You are a ",gender$
appendInfoLine: "Your result is:",result,"%"
```

As is evident from Figure 6, the script performs the required task. It uses the form control to request some information from the user. The form entries are enclosed between *form* and *endform*. The *form* keyword is followed by the name or caption given to the form. Form description can be provided using the *comment* command. All form fields have the following structure:

```
<data type> <field or variable name> <default value>
```

The various data types supported in Praat are: a *sentence* for a text input, an *integer* for a numeric integer value, *real* for a real valued input, *Boolean* for a binary value, etc. Values stored in variables can be displayed simply by typing variable names. If the variable is of a text type, then a \$ sign follows the variable name. Clicking on *Run* executes the script.

Invoking a script from another script

Assume that you want to call a script from another script. For instance, consider that you want to write a script, which, when executed, in turn calls and runs the above script (discussed in Example 1) with specified arguments. If the target external script is *formTest.praat*, then a script with the following statement will directly execute the script with user-provided arguments.

```
runScript: "formTest.praat", "Albert",35,"Male",86.57
```

Example 2: To construct a simple signal with the frequency value provided by the user, play its tone, and show its spectrum and spectrogram, type:

```
# Form begins..
form Signal Analysis
    comment Please provide the signal frequency and
type of analysis you want to perform.
    positive frequency
    choice analysis:
        button Spectrum
        button Spectrogram
endform

#Create sound with default parameters
Create Sound as pure tone: "tone", 1, 0, 0.4, 44100,
frequency, 0.2, 0.01, 0.01

#play the sound
Play

#clear the info window
clearinfo

if analysis == 1
    #Show the frequency spectrum
    To Spectrum: "yes"
    View & Edit
else
    #Show the spectrogram
    #Edit
    To Spectrogram: 0.005, 2000, 0.002,
20, "Gaussian"
    View
endif

#display the message
writeInfoLine: "You have just constructed and listened to a
",frequency,"Hz sound signal."
```

The script to solve the problem given in Example 2 has been presented here. Now, the execution of the script is left to you.

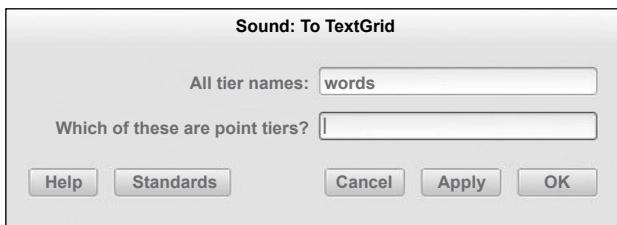


Figure 7: Creating TextGrid

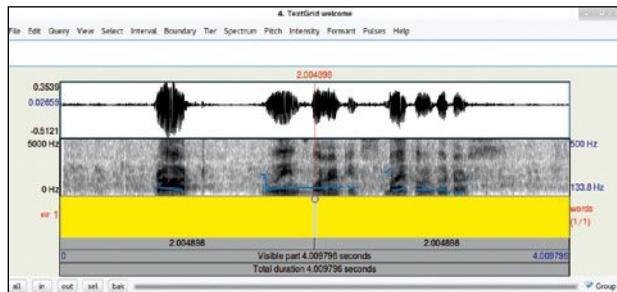


Figure 8: Annotating sound

You must have noticed that some commands expect a long list of parameters as their arguments, and it is often cumbersome to remember all such options. Praat helps here by introducing a *History* command, which keeps track of all the operations carried out during a specific session. Selecting *Edit >> Paste history* in the Praat script writer will populate the entire editor with all the commands used so far. You can keep the required commands and remove others manually, and that's it! *Edit >> Clear history* flushes out the entire history. The details for the *To spectrogram* command parameters can be availed by selecting a specific sound object and clicking on *Analyse Spectrum >> To spectrogram*.

You can play with this script by changing various parameters and options, and observe the effect on the signal.

Sound annotation

Annotating sound involves supplementing metadata to sounds, and providing additional information about the sound. Annotations are useful in building a speech corpus for linguistics research by transcribing speech.

Utterances from a sound file can be segmented into smaller units such as words, syllables or phonemes. This is done by establishing boundaries between two consecutive sound contents at different time stamps and separating them. TextGrids in Praat are used to record timestamps and create boundaries by adding multiple tiers to the sound data. There are two types of tiers:

1. Interval tiers, which are made up of time duration, characterised by starting and ending times (such as phonemes, syllables, words, etc).
2. Point (Text) tiers, that represent events which occur instantaneously at a single point of time, and are basically found in the presence of suprasegmental features (such as the pitch peak, glottal closure, etc).

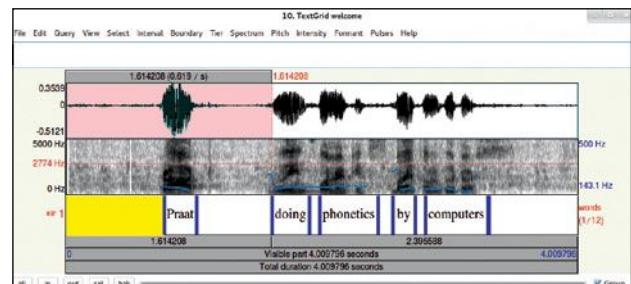


Figure 9: Annotations for given utterance

Let us create a simple annotation at the word level. Follow the steps listed below to generate a simple interval tier corresponding to words.

- Load the sound file which you need to annotate into the Praat *Objects* window, and select it.
- Click on the dynamic menu button *Annotate >> To TextGrid*. A form window will appear as shown in Figure 7. Here, you can specify the names of all the tiers you want to add in your TextGrid. In our case, we have only one tier in the TextGrid, namely, words.
- Select the sound file and the TextGrid object together by pressing the *Ctrl* button, and click on *View & Edit*. This will open the TextGrid as shown in Figure 8. You can observe that, below the sound signal, a layer has been added, where you can specify your transcription. Above the sound signal is a text area, where you can write the contents for the selected segment, highlighted in yellow.
- To determine the boundary of a segment, select a portion of the speech segment, click on the red dotted vertical line first, and then click on the small circle at the top, which will create a boundary for you. Click on the segment you wish to annotate, and add your content.

Figure 9 describes a completely annotated speech utterance at word level. 

References

- [1] Speech Signal Processing with Praat, David Weenink, <http://www.fon.hum.uva.nl/david/sspbook/sspbook.pdf>
- [2] Praat website, <https://www.praat.org>
- [3] Annotating audio files, <http://www.krisyu.org/blog/posts/2013/06/annotating-audio-files/>
- [4] Tutorial on Praat for Beginners, <http://swphonetics.com/praat/tutorials/understanding-waveforms/speech-waveforms/>
- [5] Introduction to Acoustic Phonetics and Praat, http://ec-concord.ied.edu.hk/phonetics_and_phonology/wordpress/learning_website/chapter_1_introduction_new.htm#1.2

By: Sapan H. Mankad

The author works as assistant professor at the CSE department, Institute of Technology, Nirma University. You can contact him at sapanmankad@nirmauni.ac.in or visit his website at <http://www.nirmauni.ac.in/ITNU/Faculty/Prof-Sapan-H-Mankad>



“Open source technology plays a crucial role in the cloud business”

Although the cloud has been around for a few years now, the phenomenon is still a mystery to many. CIOs still watch their step while moving on to the cloud. And when it comes to the open source cloud, it gets even more complicated for some. That is why we have scheduled CIOs from some of the country's biggest firms to talk about the cloud and open source at Open Source India 2015, on November 19 and 20 at the NIMHANS Convention Centre in Bengaluru.

In this interview, **Diksha P Gupta** from *Open Source For You* speaks to **Sunil Gupta**, **executive director and president, Netmagic**, about how open source and the cloud is changing the modern IT landscape. Excerpts:

Q Netmagic has come up with a data centre in Mumbai. What markets and sectors are you looking at with this data centre (DC)?

Typically, I divide my customers into two segments. In our terminology, we call them ICE and NICE. ICE stands for Internet centric enterprises and NICE for non-Internet centric enterprises. Under ICE, we include companies like Amazon, Flipkart, Myntra, et al. You must be aware of the growth wave of e-commerce in India. It is churning out new players every day. This is one market that Netmagic

has a good hold on. Almost all of the e-commerce companies in India are hosted with us. So, this will continue to be one important sector for us, and will take up space and services in our new data centre.

The second set of customers is obviously those who are non-Internet centric. These could be in manufacturing, banking, financial services, insurance, the entertainment industry... almost all the industries practically become candidates for us. So existing customers could scale up their operations and take up space in the new data centre,

or new customers from the above mentioned segments could do so.

Q Players like Microsoft and Amazon are now setting up DCs in India to tap sensitive sectors like the government, BFSI and telecom. With these players coming in, the competition would increase. How do you plan to face it?

It is true that players like Microsoft and Amazon deliver a fairly large scale, specially-operated, automated cloud delivery model. With them, everything is automated through a portal and customers can just go there and consume the cloud. They can use it or come out of it any time. So this is one way of offering cloud services, which involves a specially automated way of delivery where you can go and consume cloud resources right away. You may have some varieties and different types of packages for customers of different sizes or with different requirements. This model typically works on a large scale. This will be an advantage for Amazon and Microsoft. But, the only drawback is that they are not delivering these services from India, but from Singapore or Hong Kong and naturally, from a regulatory compliance point of view, many of the customers are not comfortable in using their services.

However, companies like us who deal with enterprise customers cannot adopt the same approach. Enterprises do not take their applications to the cloud straight away. Decision-makers in an enterprise like to first ask questions that will help in operating an app on the cloud. They would have concerns over their existing IT infrastructure. So, the very first problem that a CIO would like to be addressed by any service provider is the right level of absorption and accommodation of the firm's existing IT infrastructure. Also, CIOs would want to lessen their troubles by outsourcing their existing structure. This would take care of issues with power, cooling, scalability and all other such things involved in running in-house infrastructure. These issues are going to be addressed by players like us, no matter who enters the market. Enterprises with legacy IT will continue to favour players like us, as they can continue utilising their existing IT hardware and avoid the cost of a fresh investment.

So, with most of the enterprises including medium, large and very large firms, the solution is never a single view of shifting to the cloud. It is a combination of using the existing investments in IT, which they would like to do for as long as there is life in them. And even when customers want to launch new applications, they would like to move to the cloud gradually.

Q What role does open source technology play in Netmagic's scheme of things?

Open source technology has a crucial role to play, particularly in the cloud and data centre business. Customers are looking for cost-effective solutions, and open source is able to provide these. A fairly large part of our business is based on open source technology. We have about 600 members in our engineering team, and most of them work on open source technologies. We support all the customers who bring in open source technologies and host them in our data centre. We are constantly dabbling with open source technologies to possibly launch another version of our cloud, which will basically be running on open source.

Q How important do you think open source technology is for enterprise IT businesses?

It is extremely important. Modern day CIOs want freedom and don't want to lock themselves in with a particular vendor. We have seen so much of migration, with our

customers moving from closed technologies to open source technologies. This is helping them and giving them a much wider variety of vendors, as well as a broader, more varied ecosystem for support services. They are not stuck with a single technology or a single vendor. It is helping

them reduce the cost of technology and cost of operations as well. Netmagic is helping here by supporting most of these technologies. I am seeing open source technology as a huge wave from my customers' point of view.

Q What are the challenges IT heads are facing in the modern IT enterprise, and what are the problems that you are trying to address?

I will speak more with respect to my field of the data centre and the cloud. The biggest challenge an IT head faces is how to manage the firm's IT investments. It started about 15 years ago when all CIOs wanted to create their own data centres within their own buildings. Everybody loved to have that corner glass room called the server room or data room. People would put in racks and have a small data centre in every building, just to be able to say that, "I have my own captive data centre."

Over a period of time, the technologies and the densities of data centres evolved from a small rack to about a 20kVA power rack. So much of compute and storage got stuffed into a single rack footprint, and the amount of power and cooling required in a data centre

got so high that a regular commercial building just couldn't handle it. There are innumerable examples of CIOs and CTOs who have talked of this problem — about not being able to have that kind of power in their data centre. They have cooling problems they cannot handle because the businesses are scaling up, big time. Even if they want to put more racks to accommodate more servers, they don't have the space, with offices all around. These are the challenges that CIOs and CXOs are facing. And that's why there is so much of migration among customers — they are closing their captive data centres and moving to third party data centres, which are giving them the same power, cooling and scalability, all in a much better way.

The second big challenge is that the businesses are growing crazily. Companies in e-commerce are growing really fast. They place requirements in front of the IT team to launch an application practically the next day. There is also a budget constraint put in by the CFO of the company. So one ultimately ends up 'oversizing the hardware' and this is the typical standard problem for all CIOs and CTOs. The solution to this is the cloud, and this is why this big wave of cloud adoption has happened.

You can just go and purchase a cloud instance of a reasonable size, and launch the application tomorrow without bothering about sizing beyond a point, because if you find that the cloud purchased is not meeting your requirement, you can actually increase that at a click of a button. And if you think you have taken a bigger infrastructure than needed, you can actually reduce it with the click of a button. The best part is that you are not stuck with a particular technology.

Q What are the kind of questions an IT head has to ask to cloud service providers while investing in the cloud?

Essentially, the first question should be, 'Can I adopt the cloud?' Different types of customers may have different needs. There are customers who are savvy, having dabbled with this technology. Such customers do not take time to adopt it, as the cloud is also becoming a commodity now; there are so many customers in the e-commerce domain who have become so very comfortable with the cloud that they don't ask any questions, they just ask us to deliver. However, at the other end of the spectrum, there are still people in traditional enterprises like those in the manufacturing domain, who have a lot of questions. The questions do not really start with a transactional analysis of the cloud offering. They include, "Can you advise me on whether I can move to the cloud?"; "If yes, what applications can I move to the cloud?"; "If not, why can't I move to the cloud?" or "How exactly can I migrate to the cloud?" They want to see the entire way forward.

The questioning from such clients is very basic, and these are the right questions to ask before getting into the nitty gritty of the cloud.

Q What are the kind of workloads that are suitable for the cloud?

Frankly speaking, if this question had been asked five years back, my answer would have been different. But today, I would say that cloud technologies have evolved and matured so much that any type of critical workload that demands huge resources in terms of compute, storage and networks can be migrated to the cloud. In general, anything which runs on Windows or any flavour of Linux can easily be migrated to the cloud.

Q So what would you like to tell clients who are looking for cloud computing to improve the utilisation of their existing servers, but are also concerned whether the utilisation and optimisation will compromise the service level agreements?

First of all, both the technology and the cloud service provider model, as they exist today, are so very evolved. When the technology came in six or seven years back, this would have been a good topic of debate. But nowadays, when clients invest in a cloud grid (which may host multiple customers), they get far more redundancy, flexibility and scalability than they might want. The amount of redundancy, security, flexibility and scalability available with a cloud grid today is far higher compared to the standard one-on-one solution for a customer. So the service levels that are given on a cloud grid nowadays are at par and even better than what was offered on a simple dedicated hardware type of solution earlier.

Q What is the importance of multi-tenancy in the cloud paradigm?

The entire underlying basis of the cloud is multi-tenancy. Every single enterprise has hundreds of applications, for which it could end up investing in hundreds of servers, storage boxes, network boxes, et al. According to several research studies, not more than 30-35 per cent of the capacity, which is invested in, is ever utilised. And the balance 70 per cent capacity is just a waste of money. And that waste of money is not only in hardware, but also an investment in costly software licences. Besides, accordingly, you end up investing so much money in creating data centres, power, cooling, manpower, networking...and a whole lot more on everything. It's a huge waste of effort too.

What the cloud ultimately did was, instead of thousands of enterprises creating their own individual environments where there is so much wastage of resources, it created a centralised grid with a pool of resources. It cuts these resources into 'pieces' and gives it to individual customers. So when this is happening, if say, one customer is not using some resources, these are available for other customers to use. And that is where you are bringing in the optimisation. So the basic premise of the cloud is based on multi-tenancy, where a centralised service provider makes a huge investment in creating a vast pool of resources like compute, storage or networks.





Resetting the root password

Have you ever forgotten your root password? Is your system set by a vendor who didn't provide you a root password? The following process will reset it.

When the Grub menu loads, use the keyboard to highlight your OS option and press the edit key, which is usually 'e'. There will be a line which usually starts with 'linux' or 'kernel' and ends with 'single' or 'quiet'.

Navigate to the line and click the *edit* key. At its end, just add the following command:

```
init=/bin/sh
```

Now click on 'Enter' and then the key to boot up the system, which is usually 'ctrl+x' or 'b'. You will be logged in as root in 'Text mode'. The file in which the password is saved would most probably be in read-only mode. To change it to read-write mode, use the following command:

```
sh# mount -o remount ,rw /
```

Now you can change the password using the 'passwd' command:

```
sh# passwd  
Changing password for root.  
Enter new UNIX password:  
Retype new UNIX password:  
passwd: password updated successfully  
sh#
```

There you go. The password has been changed.

—*Sajin M. Prasad*,
sajinprasadkm@gmail.com



Protect your valuable files

The *chattr* attribute is used to stop accidental deletion of files and folders. You cannot delete files secured via *chattr* even if you have full permission over files. Even *root* cannot delete files that have the *chattr* command.

To use this command, follow the steps shown below:

```
#cat > test This test file  
#chmod 777 test  
#chattr +i test
```

...and then try to delete a test file:

```
#rm -fr test (not allow)  
#mv test (not allow)
```

If you want to remove permission, run the following command:

```
#chattr -i test
```

—*Anshoo Kumar Verma*,
anshooverma@expressindia.com



Find your public IP address with *dig*

Sites like *whatismyip.com* give you the Internet-facing IP address. This trick shows the outgoing IP address from the command line with the DNS *dig* utility:

```
#dig +short myip.opendns.com @resolver1.opendns.com
```

—*Sani Abhilash*,
saniabilash@gmail.com



Avoid auto updation of the '/etc/resolve.conf' file in RedHat GNU/Linux

In RedHat GNU/Linux, the */etc/resolve.conf* file is used to configure the DNS resolver library. If you edit the file manually, then after each reboot cycle or after restarting the network service this file will get overwritten. We can override this behaviour by modifying network configuration. To avoid auto updation of */etc/resolve.conf* just add the following entry in the */etc/sysconfig/network-scripts/ifcfg-eth<NUMBER>* file:

```
PEERDNS=no
```

To enable auto updation again of */etc/resolve.conf*, just remove the above entry or change its value:

"PEERDNS=yes".

Note: Replace '*ifcfg-eth<NUMBER>*' with the actual number.

—Narendra Kangalkar,
narendrakangalkar@gmail.com

Recording the shell session and replaying it

Here is a small tip that allows you to record and replay commands that you use in your shell.

The command to start recording the shell session is as follows:

```
# script -t 2> project.timing -a project.session
```

After this, all subsequent commands typed in the terminal will be recorded. To stop recording, you need to just 'Exit'.

The command to replay the recorded session is as follows:

```
# scriptreplay project.timing project.session
```

The script command does not have the 'Pause' option, but here is another way to achieve that through a Python module. First, we need to install the Python module as shown below:

```
#pip install TermRecord
```

Now, to start recording a shell session, use the following command:

```
#TermRecord -o mysession.html
```

All subsequent commands will be recorded. Finally, just 'Exit' to stop recording.

To view this, you can open the output file on any browser. This will let you play, replay and pause.

—Rajeshwaran Gopalakrishnan,
rajeshwaran@dcis.net

What shell is your Linux distribution using?

A shell is a special program that allows the user to communicate with the kernel. It accepts specific

commands and instructions, and communicates them to the kernel. There are dozens of shells for Linux and the most popular are: Bash (Bourne again shell), ksh (Korn shell), csh (c shell), ash (Almquist shell) and Dash (Debian almqiust shell).

ramesh@ubuntu:~\$echo \$SHELL

/bin/bash

Ubuntu uses the Bash shell. The dollar sign indicates the Bourne shell. Shell commands may vary from shell to shell, but Linux commands are the same for every shell though they differ from one Linux distribution to the other.

—Ramesh Mariappan,
ramesh2get@gmail.com

Use Bash to change a given string to either upper case or lower case

Let's look at the following string as an example:

Linux For You.

```
[raja@support ~]$ string="Linux For You"
```

Now, to convert it into lower case, use the following command:

```
[raja@support ~]$ echo ${string,,}
```

linux for you

Now convert the same into upper case, as follows:

```
[raja@support ~]$ echo ${string^^}
```

LINUX FOR YOU

—Raja Genupula,
genupulas@ubuntu.com

Share Your Linux Recipes!

The joy of using Linux is in finding ways to get around problems—take them head on, defeat them! We invite you to share your tips and tricks with us for publication in OSFY so that they can reach a wider audience. Your tips could be related to administration, programming, troubleshooting or general tweaking. Submit them at www.opensourceforu.com. The sender of each published tip will get a T-shirt.

DVD OF THE MONTH

Sample these Linux distros for geeks.



Manjaro Linux 15.09 Cinnamon Edition

This is a GNU/Linux distribution based on the independently developed Arch Linux by the community. It's a fast, powerful and lightweight distribution that provides access to the very latest cutting-edge software. A user-friendly installer is provided, and the system itself is designed to work straight out-of-the-box, in live mode.

BackBox Linux 4.4

Today's IT infrastructures are way too complex. BackBox allows you to simplify security in your IT infrastructure. It is the perfect security solution, providing pen-testing, incident response, computer forensics and intelligence gathering tools. The most current release of BackBox Linux includes the latest software solutions for vulnerability analysis/assessment and pen-testing. You can find the ISO image in the folder *other_isos* on the root of the DVD. The bundled ISO image is for 64-bit systems.

Proxmox VE 4.0

Proxmox VE is a complete open source server virtualisation management solution. It is a powerful and lightweight distro optimised for performance and usability. Proxmox is based on KVM virtualisation and container-based virtualisation. It manages virtual machines, storage, virtualised networks and HA clustering.

The enterprise-class features and the intuitive Web interface are designed to help you increase the use of your existing resources, while reducing hardware costs and administrating time. You can find the ISO image in the folder *other_isos* on the root of the DVD. The bundled ISO image is for 64-bit systems.

What is a live DVD?

A live CD/DVD or live disk contains a bootable operating system, the core program of any computer, which is designed to run all your programs and manage all your hardware and software.

Live CDs/DVDs have the ability to run a complete, modern OS on a computer even without secondary storage, such as a hard disk drive. The CD/DVD directly runs the OS and other applications from the DVD drive itself. Thus, a live disk allows you to try the OS before you install it without erasing or installing anything on your current system. Such disks are used to demonstrate features or try out a release. They are also used for testing hardware functionality, before actual installation. To run a live DVD, you need to boot your computer using the disk in the ROM drive. To know how to set a boot device in BIOS, please refer to the hardware documentation for your computer/laptop.



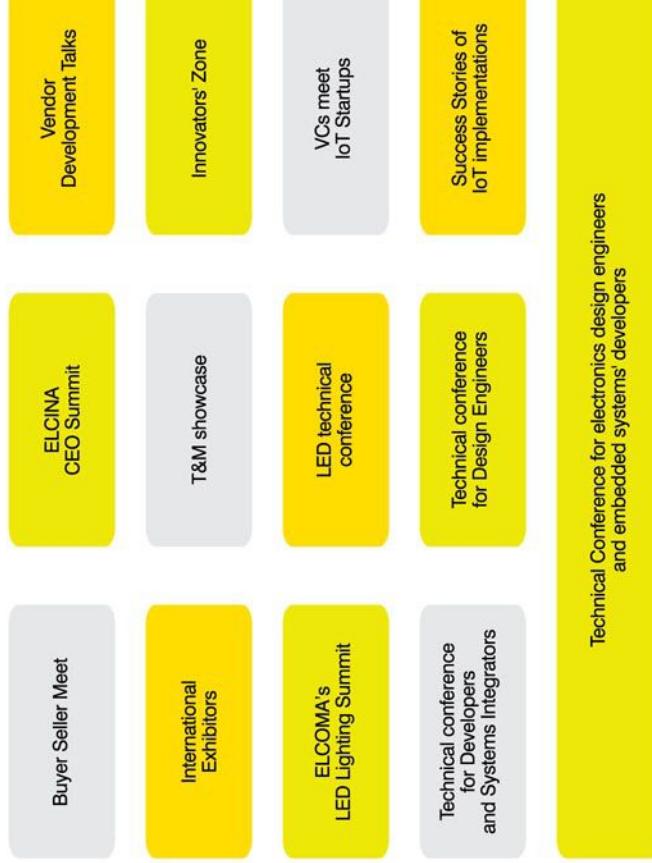
INVEST. IN. INDIA.



INDIA ELECTRONICS WEEK

January 11th-13th, 2016 • Bengaluru

What's Happening at IEW 2016



www.indiaelectronicsweek.com

An EFY GROUP EVENT

Six Events. Coming Together

Electronics For You Expo
Innovate. Design. Manufacture. Source.
An expo for OEMs to meet electronics related suppliers and service providers

electronics Rocks
India's largest confluence of technology decision makers across the electronics industry

TEST & Measurement India
Asia's leading exposition of test & measurement products and services

January 11th-13th, 2016 • Bengaluru
For more information, talk to us at +91-11-40596605 or email at growth@efy.in

January 11th-13th, 2016 • Bengaluru

LEDASIA.in
India's first event to focus on the latest technologies for LED lighting

iotshow.in
India's first large scale IoT event with an exposition and conferences

Rajeshwar India
India's first event related to the use of electronics in defence, aerospace & security

Community Partners

Gold Partner

Partner Associations

International Partners

Media Partners

Enterprise Postgres



YOU CAN MAKE IT BETTER

Discover the power of EDB Postgres to accelerate, innovate and transform, while dramatically saving costs. Join the market leaders who chose EDB Postgres for enterprise level database performance.

- GARTNER MAGIC QUADRANT LEADER
- ENTERPRISE PERFORMANCE & SECURITY
- OPEN SOURCE POWERED
- POSTGRES TRAINING & EXPERTISE
- 24/7 SUPPORT

CALL US AND FIND OUT MORE TODAY

EnterpriseDB Software India Private Limited,
Unit No.803

8th Floor 'Godrej Castlemaine'
Next to Ruby Hall

Survey No.155, CTS 43
Sassoon Road Pune - 411 001

Phone +91-20-66449600/601
www.enterprisedb.com
sales@enterprisedb.com

100%

83%

45%

33%

2

