


CPOSC 2009

Harrisburg University of Science & Technology
October 17, 2009



“When we speak of free software, we are referring to freedom, not price. Our General Public Licenses are designed to make sure that you have the freedom to distribute copies of free software (and charge for them if you wish), that you receive source code or can get it if you want it, that you can change the software or use pieces of it in new free programs, and that you know you can do these things.”

GNU GENERAL PUBLIC LICENSE
Version 3, 29 June 2007



Time	Room 1 (14 th Floor Auditorium)	Room 2 (13 th Floor - Room #1305)	Room 3 (13 th Floor - Room #1302)
8:00 - 9:00 AM	REGISTRATION / BREAKFAST (12 th Floor Hallway)		
9:00 - 9:10 AM	WELCOME (14 th Floor Auditorium)		
9:20 - 10:10 AM	The Business of Open Source Licensing (Owen Winkler)	Stop Worrying and Start Monitoring with Nagios (Andrew Libby)	Making Java Fun Again: Web App Development with Tomcat & Eclipse (Jim Capp)
10:20 - 11:10 AM	DRBD, Network Raid, High Availability and General Awesomeness (Brian Gorka)	Drupal: Challenges and Techniques for Managing Change (Christian Pearce)	The Statistical Significance of R (Premal Vora)
11:20 - 12:10 PM	MySQL Performance Tuning for non-DBAs (Tom Clark)	The Value of IRC to Open Source Projects and Groups (Jonathan Simpson)	Survey of Fuzzy String Matching Algorithms (Kyle Burton)
12:10 - 1:10 PM	LUNCH (12 th Floor)		
1:10 - 1:40 PM	Wonderful Desktop Tricks and Aesthetics (Seth Jerome)	Clojure: Usable Lisp (David Whittington)	Why You'll Love Python (Shawn Milochik)
1:50 - 2:40 PM	Virtualization is Easy Now: An Introduction to VirtualBox OSE (Bob Igo)	Proper Use of a Computing Cluster (Phil Kostenbader)	Jump Start Django: The Web Framework for Perfectionists with Deadlines (Rob Yates)
2:40 - 3:00 PM	xforty technologies REFRESHMENT BREAK (12 th Floor)		
3:00 - 3:30 PM	Watching and Manipulating Your Network Traffic (Josiah Ritchie)	Building Scalable, Fault-Tolerant Applications with Erlang (Joshua Miller)	Scratch: Programming for the Young and the Young at Heart (Michael Badger)
3:40 - 4:30 PM	Contributing to Open Source Projects (Elizabeth Krumbach)	How to Get There and What's Along the Way: Driving Directions and Geospatial Search (Eric Beyeler)	Maatkit, a Toolkit for MySQL Users (Baron Schwartz)
4:40 - 5:00 PM	CLOSING / GOODBYE (14 th Floor Auditorium)		



To access the Internet via WiFi, use "CPOSC" as both the username and password. All uppercase.

Michael Badger

Scratch: Programming for the Young and the Young at Heart

Learn why Scratch makes an ideal language to teach your children (8 and up), your students, or yourself the basics of programming by creating games, interactive stories, and multimedia projects. Scratch uses a fun, accessible environment that's as easy as dragging and dropping blocks from one part of the screen to another.

We'll create a sample project that demonstrates the "basics" of Scratch. We'll also connect a webcam and a sensory board to use real-world input for our project.

Kyle Burton

Survey of Fuzzy String Matching Algorithms

Do you like how Google suggests a fix when you misspell (I mean typo) a word? Did you ever wonder how you could implement a spell checker?

Fuzzy string matching dates back to the 1800s to the US Census Bureau's

mandate of not over-counting the US population. My personal exploration into the record linkage aspect of data integration has touched on several fuzzy string matching algorithms. Come and hear about how several of these algorithms work and some of the techniques I have used in industry settings with large datasets.

Jim Capp

Making Java Fun Again: Web App Development with Tomcat & Eclipse

Java is one of the most powerful and widespread software development



**Don't forget the
#cposc hash tag!**

platforms ever created. Apache Tomcat is a popular lightweight open source Java Servlet container and is employed by such companies as E*Trade, WalMart, and The Weather Channel. Using Tomcat, developers can leverage a universe of open source Java libraries to build enterprise quality web applications.

Developers who are accustomed to the rapid productivity afforded by scripting languages are often frustrated by Java, as it is a compiled language and is very strongly typed. Eclipse comes to their rescue by providing a development environment that removes the barbs and other sharp edges that are the result of strong typing, and offers compiling on-the-fly that makes web development in Java fun again.

This talk will explore the synergies afforded by the combination of Java, Apache Tomcat, and Eclipse, with perspectives from the development of an enterprise-wide web application written for La-Z-Boy Chair Company.

Tom Clark

MySQL Performance Tuning for non-DBAs

If you run any database driven websites, the odds are good that you use MySQL. And unless you've taken some time to configure it for your application, the odds are extremely good that MySQL is not performing as well as it can. Often, some easy configuration changes can yield big performance gains.

In this talk we will discuss techniques for evaluating your database and configuring it for better speed and efficiency.

Brian Gorka

DRBD, Network Raid, High Availability and General Awesomeness

DRBD refers to block devices designed as a building block to form high availability (HA) clusters. This is done by mirroring a whole block device via an assigned network. DRBD can be understood as network based raid-1.

I'll be covering what DRBD is, how it can be used in everyday projects, as well as integrated with clusters and VMware/Xen/KVM for DR or High Availability.

Bob Igo

Virtualization is Easy Now: An Introduction to VirtualBox OSE

I used to avoid virtualization because it looked like too much effort. When I discovered VirtualBox OSE, it was so easy to use that I abandoned physical hardware for lots of tasks. The talk will show some powerful yet simple examples of things you could do with virtual machines that are faster and easier than using real hardware. No previous experience with virtualization is necessary to attend this talk.

Seth Jerome

Wonderful Desktop Tricks, and Aesthetics

The Linux desktop has evolved by leaps and bounds over the last few years. There are many choices and variants. The most common choices today are Gnome and KDE.

Gnome has always been rock solid, and with a few enhancements, it can be quite jaw dropping. KDE has undergone major changes since the 3.0 days. With 4.0, and now 4.3, you can have a very stable, and eyecandy-filled desktop, and still maintain functional productivity.

This session will show you how to unlock your desktop and do eye catching things

with it that Windows fans will *Oooh* and *Ahhh* over... and in the end make you more productive.

Phil Kostenbader

Proper Use of a Computing Cluster

A high level presentation on the improvements in the use of cluster computing. MPI is necessary to fully utilize a cluster environment. One might think of a cluster as a group of machines unified somehow to provide a single computing facility. Yet the reality is that in many cases the user has to manually break down their calculations into many jobs, submit them to a cluster, and then assemble the resulting data. Within the past year, some software used by the scientists at Rutgers University been adapted to use MPI.

Open source used at this site includes openmpi, ganglia and torque / in addition to reporting software pbsacct, gsh and several utility scripts used to maintain user accounts, software distribution, ssh key setup, and process monitoring. Resource allocation and scheduling of jobs has taken a look similar to the early IBM mainframe days – in a sense this technology has come full circle and returned to an old fashioned batch environment.

A look at a large European scientific grid (enmr.eu), and their technique in providing access and job submission.

Elizabeth Krumbach

Contributing to Open Source Projects

Want to contribute to an Open Source project, but not sure where to begin?

This talk will seek to answer this question and cover: skills which are valuable to open source projects, how to go about getting involved, considerations to be made when getting involved, some of the expected behavior of contributors and what kinds of benefits can come from involvement.

Andrew Libby

Stop Worrying and Start Monitoring with Nagios

This talk will provide a general introduction to Nagios, it's approach to monitoring based on simple plug-ins, cover configuration principals, monitoring and notifications. This talk will be based on practical real world experiences gained from running several installs of Nagios.

Time permitting, a discussion of developing Nagios plug-ins will be covered.

Joshua Miller

Building Scalable, Fault-Tolerant Applications with Erlang

For the complete newcomer to Erlang, we'll take a look at what features have made this venerable language and platform suddenly popular. By building on functional programming principles and built-in support for lightweight, shared-nothing processes, Erlang and the OTP platform provide a great base for highly scalable, concurrent applications.

We'll take a quick tour of Erlang's occasionally idiosyncratic syntax, then dive right into a demonstration of how the message-passing model of concurrency allows us to share work across processes and machines without the problems that plague similar work in other languages.

Shawn Milochik

Why You'll Love Python

We generally don't pick our programming languages; they're chosen for us. Whatever is taught in school, or used by employers, becomes the norm. Not so with Python. The Python community is enthusiastic about the language, and is extremely active in mailing lists, conventions, and contributing to the language and the frameworks built upon it. This is because Python is an extraordinary language.

Learn why Google hired Python's creator, Guido van Rossum, made Python the only

language available for the Google App Engine upon its release, and why Python has been ported to Microsoft's .NET platform (Iron Python) and the Java Virtual Machine (Jython). Hint: It's because people really like Python, and want to use it everywhere.

This presentation exhibits Python's beauty by showing what makes it different, and how and why those differences make its users want to use Python everywhere.

Christian Pearce

Drupal: Challenges, and Techniques for Managing Change

Change is a word that Drupalers must get comfortable with. Drupals development methodology keeps the code changing constantly. Adapting to change has a lot of positive benefits: fresh technology is always available for use, bugs and security fixes are handled rapidly, and a superior codebase. It also has a lot of negative impacts: Upgrades can be difficult and are required every 18 to 24 months if you want support from the community.

Change also comes in the form of adding new features added to a website. Since Drupal is an online processing CMS, all changes need to be executed in production by hand. This sort of change ends up being cumbersome and error prone if you are not methodical about doing it on a test server first.

My talk explores the various types of change we need to deal with as a Drupal developer, how they are handled conventionally, and what sort of new development is on the horizon to help automate some of the conventional approaches.

Josiah Ritchie

Watching and Manipulating Your Network Traffic

Linux has many advanced network traffic monitoring and manipulation capabilities. In this session we'll assume you have command-line access to a Linux firewall or router between your network and the outside world. We'll explore the tools



Backchannel Chatter

irc://irc.cplug.net
http://cplug.net/irc/
Join #cposc to chat!

available to most distros and how they can be used to gather and manipulate information from the network. We'll primarily discuss tools available from the command line such as tcpdump and iproute2. The first will help us discover what is flowing and the second will help us manipulate it in the direction desired.

Baron Schwartz

Maatkit, a Toolkit for MySQL Users

Maatkit is a toolkit for MySQL users. It provides missing functionality, such as checking whether replication slaves are identical to their masters. It also has a number of convenience tools for analysis and productivity.

This talk introduces Maatkit and gives an overview of many of the tools, as well as productivity tips for using it more effectively. I'll share what we've done differently at the code and architecture level, which is fairly unique as far as I know, and has been a great success in making the tools more useful and thus more widely adopted. I'll also discuss the project's goals and development practices, which should be of interest to anyone who's involved in open-source development; we've learned a number of lessons about what works and doesn't work.



WiFi Internet access
login information:
Username = CPOSC
Password = CPOSC

Jonathan Simpson

The Value of IRC to Open Source Projects and Groups

A presentation describing how IRC channels can be beneficial to open source projects and groups. IRC channels provide a real-time method to communicate within groups, and can be an excellent tool for helping users, collaborating with developers, and planning events.

I'll talk about how to use this by creating and managing a channel on an IRC network, and how it can improve the experience of a group or project.

Premal Vora

The Statistical Significance of R

"R" is a very powerful open source package for mathematical and statistical modeling. This introductory presentation

on "R" gives the audience a short history of the origins of "R", then an overview of the kinds of problems it excels at, and finally a brief introduction to the syntax and semantics of "R".

David Whittington

Clojure: Usable Lisp

Lisp just won't die. Clojure is the latest proof of that. Come to this talk to learn why Lisp is still relevant and see a small web app developed in Clojure.

Owen Winkler

The Business of Open Source Licensing

Just because your software is open source doesn't mean it can't help you make more money. Can you gain a competitive edge by not charging for your software? How can Open Source Licensing help reduce your operational costs?

If you're already giving your software away, how can it generate income for you? Get an overview of different licenses and how you can apply them in a for-profit business environment. Learn the advantages and disadvantages of some popular licenses and what GPL, BSD, and HJTJ mean for your company.

Rob Yates

Jump Start Django: The Web Framework for Perfectionists with Deadlines

This presentation will get you up and developing with the Django web

framework quickly and easily. We will cover a high-level overview of the framework, explain how it differs from other frameworks, follow along as we build a useful website, and review the built-in testing functionality. As application reusability is a big part of the promise of Django, we will talk about how to make your apps reusable and introduce the Pinax project which bolts together contributed, reusable applications to quickly deploy a social networking site with little effort.

If you're tired of menial development tasks, can't stand writing SQL, or get frustrated at using "Enterprise" technologies for prototyping, then Django is for you.

Eric Beyeler

How to Get There and What's Along the Way: Driving Directions and Geospatial Search

Learn the behind-the-scenes details of how MapQuest performs routing and geospatial search.

As a MapQuest employee, I'll cover routing along road networks, the algorithms we use, road attributes that are considered, testing, as well as performance.

For searching, I'll discuss Lucene, spatial organization, and identification.

Help us make CPOSC even better!
Please fill out the feedback form included in your packet and return it to the registration desk!



Follow @cposc on Twitter for updates!

Platinum Sponsor



Anteil leverages technology to help its customers achieve their business goals. With over 25 years experience in data communications, Anteil has been instrumental in leading its customers in adopting, and profiting from, Internet technology. In addition to our custom software development and integration services, Anteil is uniquely positioned to help our customers merge their telephone systems with their business software applications, providing an unprecedented level of customer service, while reducing their operating expenses.

Gold Sponsors



INetU Managed Hosting provides managed dedicated hosting on the top-rated hosting network with excellent customer service and a 100% uptime SLA.



Zenoss is a leader in commercial open source network and systems management, providing solutions to monitor the availability and performance of the entire IT infrastructure.

Silver Sponsors



www.mythic.tv

We only sell HTPC hardware and accessories that are compatible with MythTV. Since 2004, we have been active members of the KnoppMyth (now LinHES) community and have applied our automation experience to enable KnoppMyth to autodetect and autoconfigure remotes, tuners, and sound devices. Unlike other stores, we primarily support Linux users! Check out our store at <http://mythic.tv>.



Virtualmin is a powerful and flexible web server control panel based on the well-known Open Source web-based systems management GUI, Webmin. Manage your virtual domains, mailboxes, databases, applications, and the entire server, from one comprehensive and friendly interface.



www.infoquest.com

Since 1994, customers worldwide have used InfoQuest Hosted Internet Business Solutions to communicate on-line, manage computers, and safeguard mission critical information. From the beginning, InfoQuest has proven to be a nimble, responsive, resilient company. Characteristics that provide our customers with innovative business solutions, dependable support and a trusted business partner. Today, just as in 1994, we're committed to the philosophy of growth as a company through the service of one customer at a time.

Bronze Sponsors





Zenoss Enterprise IT management software is designed to meet today's data center operations challenges. It is a single product that provides full stack coverage of networks, servers, applications, services, and virtualization. Functionally, it provides complete operational awareness by combining discovery and inventory, availability and performance monitoring, event management, and reporting. And since it's built with open source technology it has the reliability, cost, and support benefits that come with an active community of thousands of users.

www.zenoss.com



INetU
Managed Hosting

INetU takes web hosting out of the box with tailored hosting solutions and a dedicated Chief Hosting Officer to work side-by-side with your team. We are SAS70 Type-II and Visa-certified PCI compliant. We have worked with clients such as Microsoft, Northrop Grumman and Sure Fit Slipcovers, and have earned a 96% client approval rating. INetU realizes the importance of high performance, availability and scalability for online projects. Our network is backed by a 100% uptime SLA.

www.inetu.net



www.infoquest.com
1-800-622-4403
sales@infoquest.com

Free Fedora Core Virtual Private Server



Get 1 Month Free When You
Sign Up For Any Monthly
Fedora Core VPS.

Sign Up Online:
www.infoquest.com/linuxvpsplans



Prices start as low as \$9.95 per month.

InfoQuest Technologies, Inc. also provides

Co-Location / Dedicated Servers

Web Hosting

Windows Virtual Private Servers

Hosted Exchange 2007

SharePoint

fedora



We only sell HTPC hardware and accessories that are compatible with MythTV. Since 2004, we have been active members of the KnoppMyth (now LinHES) community and have applied our automation experience to enable KnoppMyth to autodetect and autoconfigure remotes, tuners, and sound devices. Unlike other stores, we primarily support Linux users!

Check out our store at:

<http://mythic.tv>



Virtualmin is a powerful and flexible web server control panel based on the well-known Open Source web-based systems management GUI, Webmin. Manage your virtual domains, mailboxes, databases, applications, and the entire server, from one comprehensive and friendly interface.

www.virtualmin.com

Acknowledgements

Friends of CPOSC

Mark Wilkins
Numinix Web Development

Harrisburg University

Michelle Mafnas
Ryan Riley

Web Hosting

Matt Daubenspeck

Registration

Shana Andreychek
Jaime Mulligan

Local User Groups That Rock

Central PA Linux User Group
Central PA Ruby Meetup
Western PA Linux User Group
Philadelphia Area Linux User Group



*The freedom of
Open-Source
&
rock-solid
Enterprise support*



Anteil

**A Proud Sponsor of
CPOSC 2009**

www.anteil.com
877 OS-LINUX