Multi-Vendor Advanced Unix Data Tools and Techniques

Learn via: **Classroom**

Duration: **4 Days**

Overview

This course provides insight into the advanced topics for handling and manipulating data streams on UNIX/Linux machines. We will start with an advanced view of using the shell efficiently for data manipulation. Over two days will be devoted to exploring Regular Expressions, as well as in-depth analysis of the advanced sed and awk features.

A full day is dedicated to an introduction to the Perl programming, which is such an important and portable programming environment these days.

Target Audience:

The course is aimed primarily at the accomplished system user, familiar with scripting, who needs to enhance their knowledge and ability of efficient data manipulation techniques.

Prerequisites

- Delegates must have good experience of using UNIX systems, and ideally of developing automated solutions using shell scripting. Familiarity with basic data manipulation tools (e.g. grep), basic regular expressions, and principles of scripting is essential.
- Typically, the prerequisites can be met by attending the Multi-Vendor UNIX Fundamentals course followed by the Mastering Bash Scripts course.

What You Will Learn

At the end of this course you will be able to:

- Improve the efficiency of shell scripts
- Be fully conversant with the advanced Regular Expression features
- Use advanced features of sed, such as tagging, grouping, looping , etc
- Understand the purpose and the flow of operation of awk
- Use awk's programming features, including decision making constructs, loops, etc
- Appreciate the power of Perl in platform-independent environments
- Use Perl's basic 'magic' variable types, characters, program structure and flow control constructs in simple administrative and reporting scripts
- Understand the differences in the three main tools discussed, and make an informed choice of which tool is the most relevant for the task in hand

Outline

Chapter 1: Course Introduction

- Files and file descriptors
- Methods for reading data
- Scripts and handling hints
- Locales

Chapter 2: Shell tips & trick

- Data extraction: IFS, set, and read
- Sub-string handling
- Process substitution
- Report generation and print

Chapter 3: Regular Expressions & grep

- Basic Regular Expressions (BREs)
- The grep family
- BRE quantifiers and anchors
- Grouping and back-references

Chapter 4: Advanced Regular Expressions

- Extended Regular Expressions (EREs)
- ERE quantifiers and anchors
- Alternation
- Character classes
- Replacement meta-characters

Chapter 5: Introduction to sed

- Running sed
- sed Regular Expressions
- sed line addressing
- Grouping commands
- The substitute command, and others

Chapter 6: More sed

- Data validation
- Using the hold space
- Multi-line matching

Chapter 7: Advanced Features of sed

- Flow control
- Loops and 'case' constructs
- Using shell variables
- File I/O
- Miscellaneous commands

Chapter 8: Introduction to awk

- Basic awk syntax
- awk line addressing
- awk variables
- awk regular expressions
- Simple awk commands

Chapter 9: Making awk Work

- Programming techniques
- Relational and logical operators
- String functions
- Arithmetic operators and functions

Chapter 10: Flow Control in awk

- if...else
- Loops
- Indexed and associative arrays

Chapter 11: Advanced awk

- Running awk scripts
- Complex input and output
- User written functions
- Running external programs from awk

Chapter 12: Introduction to Perl

- Running a perl script
- Comparing awk and Perl
- Perl variables
- Interpolation

Chapter 13: Perl Expressions and Operators

- Scalar operators
- Context and type conversions
- Useful scalar functions
- Array and hash functions

Chapter 14: Decision Making in Perl

- Basic conditional statements
- File condition functions

- Perl Regular Expression syntax
- Perl loops
- The ARGV filehandle

Chapter 15: Advanced Features of Perl

- Writing to files
- Subroutines
- Modules
- Advanced Perl Regular Expressions