**COP 5611 - Advanced Operating Systems**

**Weekly Project Report - 1**

**Harish Chetty (hmc13f) (Team: Diablo)**

1] Started learning Perl.

Using the O’Reilly Learning Perl Book:

Learnt about

* Scalar Data,
* Arrays and List Data &
* Control Structures

2] Read the following papers:

**File classiﬁcation in self-\* storage systems**

Michael Mesnier, Eno Thereska, Gregory R. Ganger Carnegie Mellon University, Daniel Ellard, Margo Seltzer Harvard University

**About the Paper:**

The focus of the paper is to predict the properties of a new file which will be created. They use decision trees to attain this to model these associations providing prediction accuracies that exceed 90% correct hits. This could lead to development of more accurate and efficient storage policies. The paper does work on new files by using decision trees, which could be a possible tactic to go about in our project.

**Early experiences on the journey towards self-\* storage**

Michael Abd-El-Malek, William V. Courtright II, Chuck Cranor, Gregory R. Ganger, James Hendricks, Andrew J. Klosterman, Michael Mesnier, Manish Prasad, Brandon Salmon, Raja R. Sambasivan, Shafeeq Sinnamohideen, John D. Strunk, Eno Thereska, Matthew Wachs, Jay J. Wylie∗Carnegie Mellon University, ∗HP Labs, Palo Alto, CA

**About the Paper:** The paper talks about self-\* systems which are self-configuring, self-healing, self-tuning and in general self-managing. The researchers try to tune performance and projection of storage management on their self-\* system Ursa Minor. They were able to attain sufficiently accurate predictions within 10-20% with overhead of about 6%. Thus the logic could help in reducing administration tasks. The main goal of our project is to generate patterns which could help in predicting and the focus of the paper does help us understand some techniques.

**COP 5611 - Advanced Operating Systems**

**Weekly Project Report - 1**

**Gaurav Sinha (gs13h) (Team: Diablo)**

The project progress for the week was categorized into two parts:

**1)** Learning the scripting language Perl.

**2)** Reading and analyzing the reading material associated with the project.

**1)** We were given a choice of implementing our project in either Perl or Python. We decided on Perl. As I do not have any background over the scripting language, I started to learn Perl using tutorials over the internet. The source I am using is <http://learn.perl.org/books/beginning-perl/>. I have covered the first three chapters including the basic fundamentals, using simple values and using hashes and lists.

**2)** I read through two of the papers that are associated with our project.

**a.** File classification in self-\* storage systems

*Michael Mesnier, Eno Thereska, Gregory R. Ganger, Carnegie Mellon University*

*Daniel Ellard, Margo Seltzer, Harvard University*

The paper discusses about how computer systems could classify file attributes automatically and then predict new file attributes when they get created by using strong associations between the attributes of the files and properties (like names) dedicated to it. The major tools implemented are decision tree classifiers to result in predictions with high accuracies. These predictions are further used to choose storage policies.

**b.** Early experiences on the journey towards self-\* storage

*Michael Abd-El-Malek, William V. Courtright II, Chuck Cranor, Gregory R. Ganger, James Hendricks, Andrew J. Klosterman, Michael Mesnier, Manish Prasad, Brandon Salmon, Raja R. Sambasivan, Shafeeq Sinnamohideen, John D. Strunk, Eno Thereska, Matthew Wachs, Jay J. Wylie, Carnegie Mellon University, HP Labs, Palo Alto, CA*

This paper discusses Ursa Minor which is large scale storage framework being developed at CMU. The motive of the project was to move towards self-\* ideality. The authors present experiences in dealing with performance tuning and projection as particular aspects to storage management. The authors find that the major performance predictions are accurate enough and there exists less performance overhead.