0.1 Clive

Clive is an open-source operating system written in Go by the Laboratorio De Sistemas at Universidad Rey Juan Carlos in Madrid[**hid-sp18-419-www-clive-lsub**]. The design goal is to create an environment where applications and services can be compiled along with libraries that permit them to run on bare hardware without a software stack[**hid-sp18-419-www-clive-lsub2014**]. The design is based on Plan 9, a research system developed at Bell Labs in the late 1980s and first released in 1992[**hid-sp18-419-www-about-plan9**].

0.2 HCatalog

HCatalog, which was originally known as Howl, is a component shipped with Hive that manages storage and tables. Its purpose is to simplify data storage and retrieval by providing a shared schema and data type mechanism between Hive, Pig, and MapReduce and the formats in which a Hadoop serializer-deserializer can be written (ORC, RCFile, CSV, JSON, and SequenceFile. Custom formats can be added as well. A REST API called WebHCat (originally Templeton) is also available[**hid-sp18-419-www-hc-wiki**].

0.3 Neptune

Neptune is a graph database service that was announced at the AWS Re:INVENT conference in November of 2017[**hid-sp18-419-www-tc_neptune**]. Graph databases are NoSQL databases that used graph structures to organize data[**hid-sp18-419-www-tp-graph-db**]. They are commonly used for social networking applications, but can be used for recommendation engines, logistics, and other applications. Amazon offers Neptune as a fully managed product. It supports Gremlin and SPARQL open source graph APIs. One can choose Apache TinkerPop Gremlin or the W3C standard Resource Description Framework model[**hid-sp18-419-www-aws-neptune**].

0.4 OpenDaylight

OpenDaylight is an open source Software Defined Network (SDN) controller[**hid-sp18-419-www-opendaylight**]. SDN separates network control logic from physical networking equipment. The result is that networking equipment is programmable like other computing platforms. SDN facilitates Network Functions Virtualization, allowing virtual network services (switching, virtualized appliances, and virtualized applications) to be deployed without having to deploy specialized physical devices[**hid-sp18-419-www-cio-sdn-nfv**]. The OpenDaylight project was founded by a group of large tech companies, including Cisco, Citrix, Ericsson, HP, IBM, Microsoft, NEC, Red Hat and VMware. Microsoft and VMware have since left the project[**hid-sp18-419-www-sdx-odl**].

0.5 Pig

Pig is a part of the Apache Hadoop ecosystem consisting of a scripting language called Pig Latin and a compiler that produces Map-Reduce programs. It was initially developed in 2006 at Yahoo! and taken over by Apache in 2007[**hid-sp18-419-die2015datascience**]. Pig Latin allows developers to code multiple interrelated data transformations as data flow sequences, with the goal fo making the code readable and easy to maintain. Pig optimized Pig Latin automatically and users can extend the language with purpose-written functions[**hid-sp18-419-www-pig**]. There is some overlap in

functionality between Pig and Hive, an SQL-like language that is also a part of the Apache Hadoop ecosystem. Pig tends to be favored by programmers and researchers, whereas Hive is preferred by data analysts[**hid-sp18-419-www-dezyre-pig**].

References