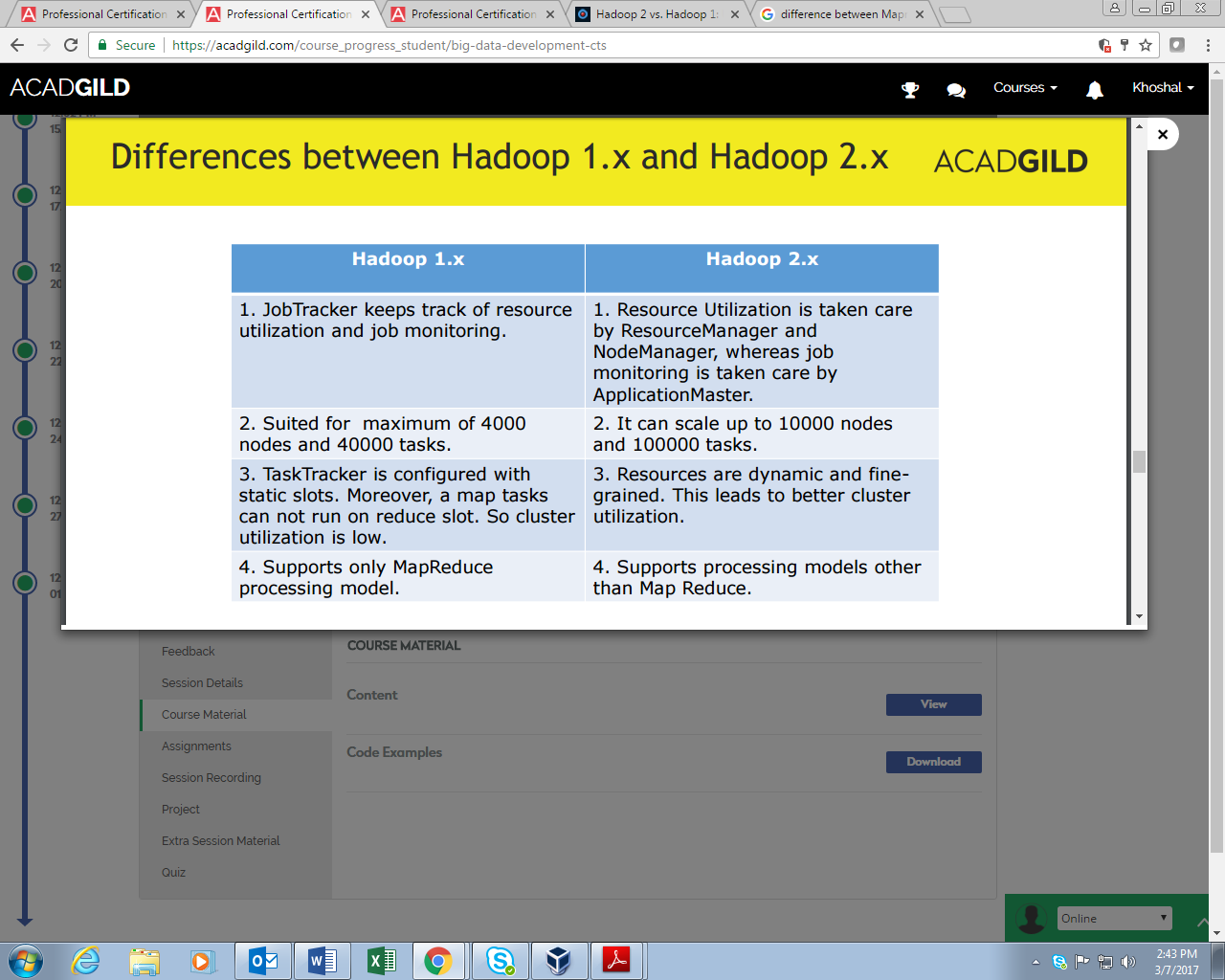
**4. What are the limitations of Hadoop 1.x and how they were overcome in Hadoop 2.x?**

ANS: The limitations in Hadoop 1.x

The major differences between Hadoop 1.x and 2.x is listed below.

In Hadoop 1.x, a single Namenode manages the entire namespace for a Hadoop cluster. With HDFS federation, multiple Namenode servers manage namespaces and this allows for horizontal scaling, performance improvements and many more name spaces.

The HDFS federation allows existing Namenode configurations to run without changes. For Hadoop administrators, moving to HDFS federation requires formatting Namenodes.



**In MapReduce1** (also called as Hadoop 1) where the HDFS Resource Management and Scheduling) and MapReduce are tightly couple. Because of this non-batch applications cannot be run on the Hadoop 1. It has single namenode so, it doesn’t provide High Availability and scalability.

**In MapReduce2** (also called Hadoop 2) where the resource management and scheduling tasks are separated from MapReduce which is separated by YARN and scheduling layers lies beneath the MapReduce layer. It also provides high system availability and scalability as we can create redundant NameNodes. The latest feature of snapshot through which can take backup of filesystems which helps disaster recovery.

The features like **“HIGH AVAILABILITY”** and **“SCALABILITY”** were missing in the Hadoop 1.x that are overcome in Hadoop 2.x.