Nepal college of information echnology

Assessment

|  |  |  |  |
| --- | --- | --- | --- |
| Level: Bachelor | Semester – Fall | Year : 2014 | |
| Programme: BE | | Full Marks : 100 | |
| Course: Big Data | | Time : 3hrs. | |
| *Candidates are required to give their answers in their own words as far as practicable.* | | |
| *The figures in the margin indicate full marks.* | | |
| Attempt all the questions. | | |

|  |  |  |
| --- | --- | --- |
| 1. a. | Explain the term "Big Data". How can you say that your organization suffers from Big Data problem. | 7 |
| 1. b. | Explain the role of distributed system in Big Data. You can provide illustrations with your case study or example if you like. | 8 |
| 2. a. | Explain the implications of "Big Data" in the current renaissance of computing. | 7 |
| 2. a. | With diagram, explain general architecture of GFS | 8 |
| 3. a. | Explain the control flow of write mutation with diagram | 7 |
| 3. b. | Why do we have single master in GFS managing millions of chunk servers? What are done to manage it without overloading single master? | 8 |
| 4. a. | Explain term 'NoSQL'. | 7 |
| 4. b. | Why does normalization fail in data analytics scenario? | 8 |
| 5. a. | Explain CAP theorem and its implications on Distributed Databases like NoSQL. | 7 |
| 5. b. | Explain Elastic Search as a search engine technology. | 8 |
| 6. a. | How would you use Elastic Search for implementing search engine in your project requiring search facility. | 7 |
| 6. b. | List components in a basic MapReduce job. Explain with diagram how does the data flow through Hadoop MapReduce. | 8 |
| 7. | Below is a data schema of three files used to detect possible crime that can happen on citizens. Write map/reduce program that finds list of users with same age. The output must be of the form (age)=>{list of userid's with age} | 10 |

