Prof. Dr. Stefan Tai   Your name: _____

Marco Peise

Enterprise Computing (WS 2016/17)

Exercise 4 (3 Portfoliopunkte)

Info:

- The solution to this exercise must be handed in by Wednesday, Nov 30th 2016, 12AM.
- Any written solution must be in an accessible PDF. Any source code in a separate ZIP File. All Files are uploaded in the Information System for Instructors and Students in a single ZIP File (https://isis.tu-berlin.de/course/view.php?id=8586).
- Please write your name on the solution sheet.

Task 1 – Misc. (10+10+10 = 30%)

a) Explain with your own words how distributed storage engines (like Dynamo) is handling temporary failures and give an example.

b) Give a reasons why one should use distributed caching and explain it in a use case.

c) Name three advantages of distributed caches and explain them in one sentence.

Task 2 – GFS (20+10+10 = 40%)

a) Which operations are supported by GFS? For which types of queries is GFS optimized?

b) State two main differences between Dynamo and GFS with regard to operation and data types?

c) Explain why GFS is particularly well suited as storage system for MapReduce job input and output data (why not Dynamo or a relational database?).

Task 3 –BigTable (20+10=30%)

a) Please describe in your own words how the BigTable storage engine works by explaining the steps of a write request, and a read request, respectively (local persistence).

b) True or False: Bigtable provides additional security, geo-distribution and support for big files. If False, shortly explain.