

Chaos Group Homework Assignment

Streaming Sort Service and Client using .NET

1. Implement a network service (using WCF) which performs the following functions:

- Allows clients to stream a sequence of strings – text – without having to specify the length of the stream in advance
- Produces a sorted sequence from the input stream for each session with a client
- Responds to a client's request to retrieve the currently streamed/accumulated data in its sorted version

As a summary, the service must support the following APIs:

- `BeginStream()` : begins a client session and returns a guid (globally unique identifier) for the stream
- `PutStreamData(streamGuid, string[] text)` : sends some text upstream; clients can call this API multiple times for the length of a session
- `GetSortedStream(streamGuid)` : streams back the data accumulated so far, sorted ascending lexicographically, treating each line of text as a single word
- `EndStream(streamGuid)` : ends a client session and closes the stream; subsequent calls to retrieve the stream data must fail.

Please note: new data can be streamed at any moment unless EndStream has been called.

The service must support multiple client streams simultaneously.

2. Implement a console application client which generates text data on the fly and streams it to the service implemented in 1.

The console application must verify that the service has performed its tasks correctly.

3.* Extra task if the candidate has time – not a requirement when considering the assignment

Implement a console application which sorts large (> 16GB) text files using the network service implemented in 1.

The task will be judged in terms of:

- Coverage of the functional requirements – does it implement what was asked for?
- Code quality – does the code share a common style? Is it well documented? Is the functionality broken down into readable/manageable chunks? Etc.
- Performance – does the code perform its tasks in the most efficient way?
- Architecture – how did the author approach the given problem set? What choices of algorithms and data structures were made? How is the code structured logically?

Expected time to complete the assignment is 1 week after acknowledging that it has been received.