

Theme : Big Data & Network

Subject : Data Reduction during Network Transfer using New Deduplication/Compression Method

Introduction

The goal of this research project is to explore new Deduplication/Compression Method that can reduce Network Bandwidth Usage and increase Data Transfer Speed. As for the Topic of Deduplication/Compression, It has been Surveyed and Implemented for long time. However, As Mobile and Cloud is prevailed, it has so much importance in data reduction more than ever.

In case of new Deduplication/Compression Method, we expect Light-Weight Deduplication/Compression Method that can be ported to embedded devices such as smart phone and sensor nodes

Scope

Challenges in Light-Weight Deduplication/Compression technologies include:

- High-Speed Deduplication Algorithm that can detect duplicate data pattern with chunk sizes of under 100 bytes
- Deduplication/Compression Algorithm that require less CPU and Memory compared to Conventional Deduplication/Compression Method
- High Speed Data Transfer Method that uses Light-Weight Deduplication/Compression
- Deduplication Data types include most of the data Type such as text data, binary data, even compress video data

Research questions

We are interested in the following research questions. These questions are not exhaustive but different research questions are open to discuss with research partners.

- What would be the method of Deduplication that requires less CPU/Less Memory/Less I/O ?
- What would be the High-Speed Data Transfer Method that uses Deduplication/Compression and Transport Optimization both ?
- Is it possible to realize Deduplication algorithm model that does not use index/chunk dictionary?
- Is it possible to port to constrained embedded devices such as smart phone or sensor nodes?

Expected Deliverables

The following is open to discussion:

- Suggestion of new Deduplication/Compression
- Suggestion of High-Speed Data Transfer method that use Transport Optimization and Deduplication/Compression
- Detailed progress reports every 3 months summarizing accomplishments.
- Prototype samples
- Patents with Samsung SDS (if agreed)