Multi-Sourced Data Retrieval in Groomed Elastic Optical Networks

Juzi Zhao and Vinod M. Vokkarane
Juzi_Zhao@uml.edu and Vinod_Vokkarane@uml.edu

Department of Electrical and Computer Engineering
University of Massachusetts Lowell

2017 International Conference on Optical Network Design and Modeling (ONDM)
@ Budapest, Hungary, May 15, 2017

CARGONET project CNS-1406370
CCDNI project CI-1541434

PROPER project DESC0012115TDD
Outline

✓ Introduction
✓ Problem Statement
✓ Proposed Heuristic
✓ Numerical Results
✓ Conclusions and Future Work
A promising solution for next-generation high-speed optical transport that provides higher levels of flexibility and efficiency to the spectral domain

- **Flexible subcarrier allocation**
- **Bit rate - different modulation level based on transmission reach limit**

<table>
<thead>
<tr>
<th>Request</th>
<th>Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>R₁</td>
<td>14 Gbps</td>
</tr>
<tr>
<td>R₂</td>
<td>5 Gbps</td>
</tr>
<tr>
<td>R₃</td>
<td>7 Gbps</td>
</tr>
</tbody>
</table>

Fixed-grid WDM 10 Gbps

Multi-Sourced Data Retrieval

- Data is replicated at many locations in Today’s data center networks

- **Traditional data retrieval**: a user has the choice between several storage sites, but can at most pick one (Single-Sourced)

- **Multi-sourced data retrieval**: parallel data transfer from several repositories to one destination

- Large datasets can be efficiently transported over multiple paths from multiple replicas
Multi-Sourced Data Retrieval

Retrieval request:

\[ R = (\{S_1, S_2, \ldots, S_H\}, d, F) \]

- Destination: \( d \)
- Data replicas at \( H \) nodes
- Data is divided into \( F \) segments

\[ F = 7 \]
\[ H = 4 \]

J. Zhao and V. M. Vokkarane, “Reverse manycast data retrieval in Elastic Optical Networks”, ICNC'17
Erasure-coded Data Retrieval

✓ Erasure coding to increase reliability (Azure and Google)

✓ \((n, m)\) code \((m \leq n)\): data is encoded and stored in \(n\) storage nodes such that the pieces stored in any \(m\) of these \(n\) nodes suffice to recover the entire data

- 4 pieces
  - A
  - B
  - C
  - D

  code rate \((4, 2)\)

✓ Erasure-coded data retrieval from distributed repositories to a single site

TCP

✔ TCP with coding
  - Google's Quic (Quick UDP Internet Connections)
  - Support a set of multiplexed connections between two endpoints

✔ Multi-path TCP with single sender

✔ TCP for Multi-source (multi-destination) connections for coded data
Multi-Source TCP

✓ Network coding based multi-sourced content retrieval service

Traffic Grooming Architecture in EONs

Single-Hop Requests $R_1$, $R_2$, and $R_4$
Multi-Hop Request $R_3$
### Scheduling - without grooming

$R_1(A, B, t_2, t_4)$, $R_2(A, B, t_3, t_5)$, $R_3(A, C, t_4, t_6)$, $R_4(B, C, t_1, t_6)$

<table>
<thead>
<tr>
<th></th>
<th>$t_1$</th>
<th>$t_2$</th>
<th>$t_3$</th>
<th>$t_4$</th>
<th>$t_5$</th>
<th>$t_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_4$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_6$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_8$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_9$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$t_1$</th>
<th>$t_2$</th>
<th>$t_3$</th>
<th>$t_4$</th>
<th>$t_5$</th>
<th>$t_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_4$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_6$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_8$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_9$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Scheduling – with Single-hop grooming

\[ R_1 (A, B, t_2, t_4), \quad R_2 (A, B, t_3, t_5), \quad R_3 (A, C, t_4, t_6), \quad R_4 (B, C, t_1, t_6) \]
Scheduling – with Multi-hop grooming

R¹ (A, B, t₂, t₄), R² (A, B, t₃, t₅), R³ (A, C, t₄, t₆), R⁴ (B, C, t₁, t₆)
EON Node Model

- Each node has a fixed number of transponders
- Each link has a fixed number of subcarriers
- Transmission reach limit for each modulation format
- K pre-computed shortest distance paths for each pair of nodes
- Guardband between subcarrier bands
- Spectrum continuity and contiguity constraint
- Time-slotted system
Erasure-coded Multi-sourced Data Retrieval in EONs

Request:
- a destination node
- code rate \((n, m)\): select \(m\) out of \(n\) candidate source nodes
- \(F\) data segments on each source node
- deadline requirement

code rate \((4, 2)\): \(n=4, m=2; F=3\)

\[
\begin{array}{cccc}
A & B & C & D \\
A_1 & A_2 & A_3 & B_1 & B_2 & B_3 & C_1 & C_2 & C_3 & D_1 & D_2 & D_3 \\
\end{array}
\]

↑
data segment
Problem Statement

✓ Traffic: Dynamic requests with time deadline
✓ Objective: minimize the request blocking
✓ Scheduling:

for each request:
   Select $m$ nodes from $n$ candidate source nodes

for each data segment transfer (from selected source)
✓ transmission starting time slot
✓ path
✓ subcarriers
✓ modulation format
✓ transponder resources (with or without grooming)
Problem Statement

- Meet the deadline of each request
- Efficiently utilize the subcarriers and transponders

$n=4$, $m=2$; $F=3$
$s_1$ & $s_4$ selected

candidate storage sources

$A_1$ $A_2$ $A_3$ $B_1$ $B_2$ $B_3$ $C_1$ $C_2$ $C_3$ $D_1$ $D_2$ $D_3$
Grooming-enabled Minimum Resource Algorithm (GMinR)

✓ **Step 1**: Select $m$ of $n$ candidate source nodes based on transponder and subcarrier resources at each node.

✓ **Step 2**: Schedule the transmission of each data segment one by one using a weighted auxiliary graph.
  - Minimum resources (transponder, subcarriers) that satisfies the deadline is allocated to the data segment.
  - Considering existing lightpaths for grooming.
  - Dijkstra’s minimum weight end-to-end path used.

✓ **Step 3**: If any data segment transmission exceeds the deadline, block the entire request.
Multi-source Data Retrieval Policies

✓ **Baseline**
  - multi-source coded retrieval with no traffic grooming
  - New lightpath for each data segment transmission

✓ **SGMinR**: Source Grooming-enabled Minimum Resource Algorithm
  - Only consider existing lightpaths with the same source in the auxiliary graph

✓ **GMinR**: Grooming-enabled Minimum Resource Algorithm
  - Both source and intermediate grooming is applied
  - Consider grooming on existing lightpaths at any hop in the auxiliary graph
# Simulation Assumptions

## 14-node NSF network

<table>
<thead>
<tr>
<th>K paths</th>
<th>3</th>
<th>Transmission reach (TR) for BPSK</th>
<th>5000km</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guardband</td>
<td>1</td>
<td>TR for QPSK</td>
<td>2500km</td>
</tr>
<tr>
<td>Simulation seeds</td>
<td>30</td>
<td>TR for 8QAM</td>
<td>1250km</td>
</tr>
<tr>
<td>Source-Destination pairs</td>
<td>random</td>
<td>TR for 16QAM</td>
<td>625km</td>
</tr>
<tr>
<td>Transponder capacity</td>
<td>400Gbps</td>
<td>Number of requests</td>
<td>10,000</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.5</td>
<td>Sub-carriers per link</td>
<td>320</td>
</tr>
<tr>
<td>Data segments/request (F)</td>
<td>uniform [3, 5]</td>
<td>Erasure coding</td>
<td>(9, 6)</td>
</tr>
<tr>
<td>Data segment size</td>
<td>15,000 Gb</td>
<td>Deadline</td>
<td>50 time slots</td>
</tr>
</tbody>
</table>
Multi-source data retrieval saves 37% in completion time compared with single-source policy.

J. Zhao and V. M. Vokkarane, “Reverse manycast data retrieval in Elastic Optical Networks”, ICNC'17
Dynamic Multi-Source Retrieval Grooming

30 new requests per time slot

- GMinR-70 better than Baseline-90
- SGMinR-60 better than Baseline-65
- Average Logical Hops: 1.132-1.093

50 transponders per node

- GMinR reduces blocking by up to a factor of 3
- Average Logical Hops: 1.091-1.184
Conclusions

✓ We investigated the dynamic multi-sourced erasure-coded data retrieval problem in groomed elastic optical networks

✓ We proposed both multi-hop grooming (GMinR) and single-hop grooming algorithm (SGMinR)

✓ GMinR can reduce 30% request blocking and 22% transponders compared with the case without grooming

✓ Future work:
  ▪ Sliceable/Multi-flow Bandwidth Variable Transponders
  ▪ Batch request allocation
Thank you

Questions?

Vinod M. Vokkarane
Vinod_Vokkarane@uml.edu