OPTIMAL FILE SIZE DISTRIBUTION AND DATA AVAILABILITY USING CFDP

EDHPC 2025

G. B. DE GIORGI, C. LEGENDRE, P. BIHN

PATRICK BIHN, 13.10.2025

MOTIVATION

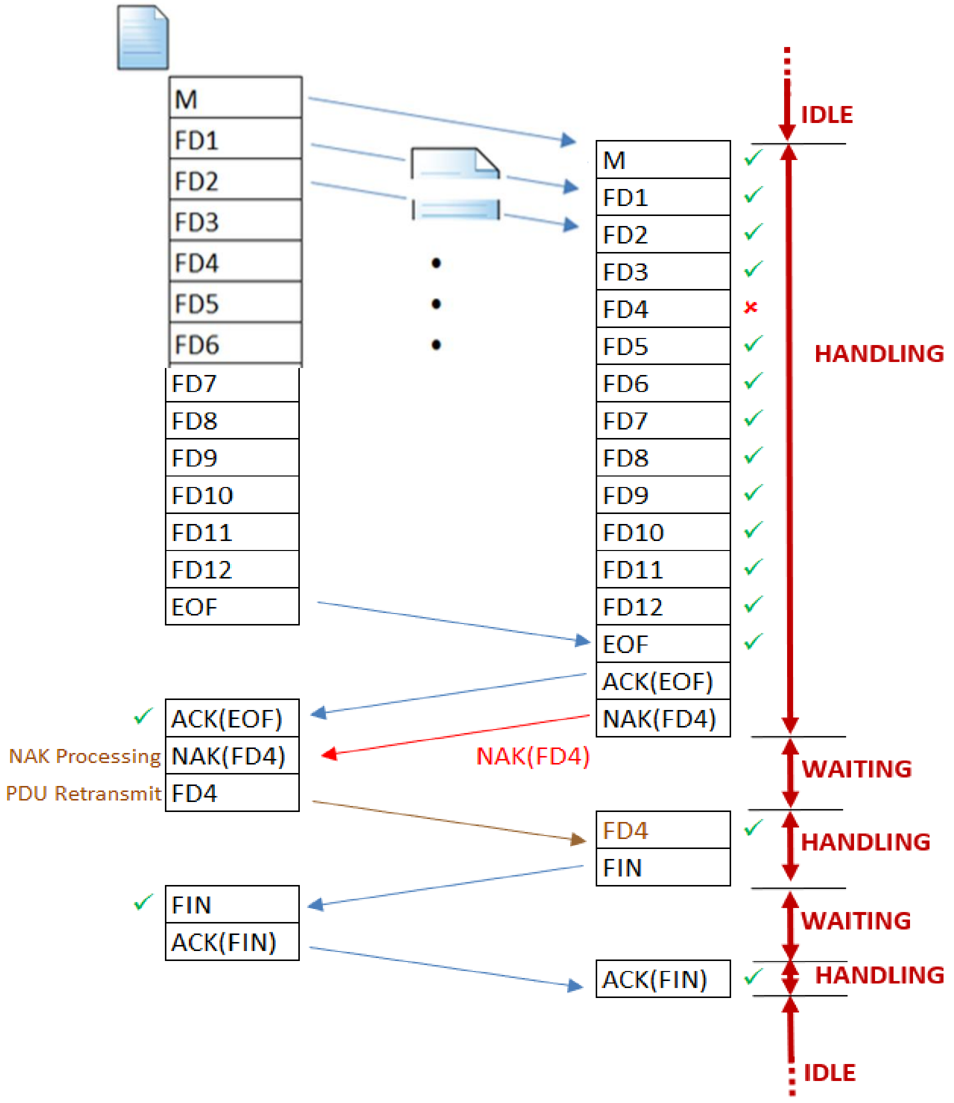
WHY SIZE MATTERS AND WHAT THIS WORK ADDRESSES

- Future Missions
  - Higher Downlink Data rates & larger data volumes
  - Earth orbiting mission with short ground contact duration (typically ~10 - 15 min per pass)
    - Links are faster, but less reliable
- File-Based Operations (FBO) + CCSDS File Delivery Protocol (CFDP) adopted to improve data handling
- File size directly impacts:
  - Class 1: data availability under disruptions
  - Class 2: number of open transactions & memory limits

CFDP OVERVIEW

CFDP IN A NUTSHELL

- Class 2 (closed loop):
  - ACK/NAK and retransmission supported
  - Requires uplink for directives
- Secures Data Link via
  - Positive Acknowledgment
  - Negative Acknowledgment
  - And retransmission of lost file portions
- File integrity ensured (if contact window allows)
- Latency and number of open transactions become critical.

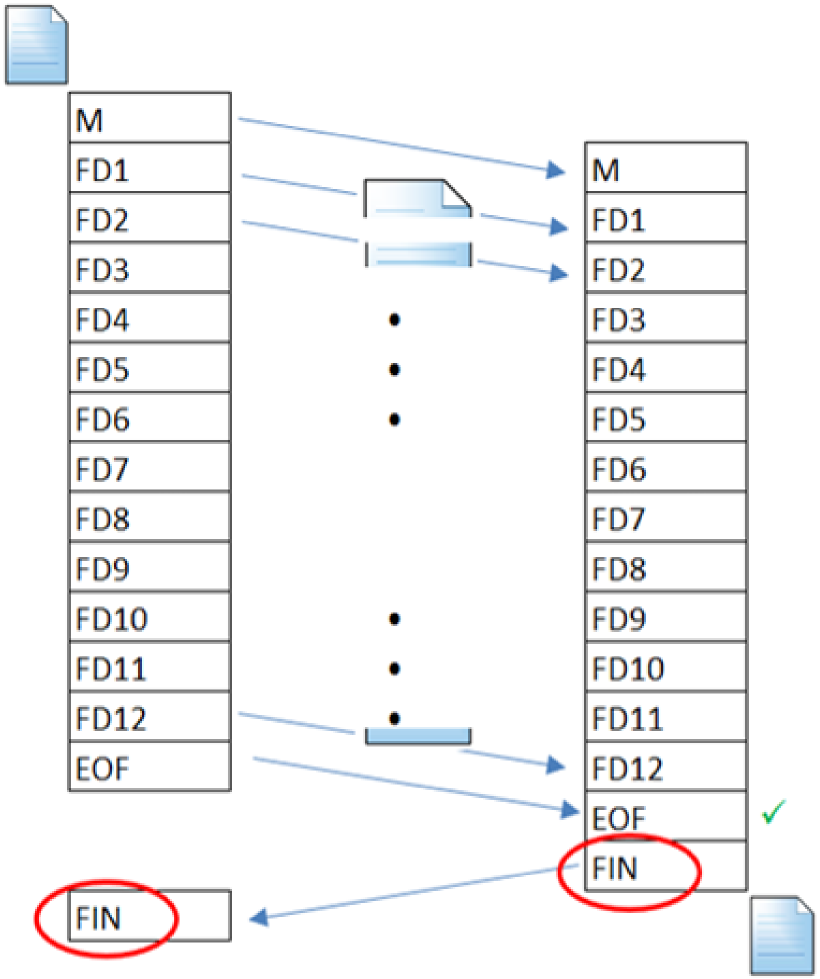


PDU -> Protocol Data Units
Metadata Attached data transferred with file data itself. M
File data A given portion (Nth) of the file data to transfer. FD(N)
EOF The End Of File notification (sender to receiver). EOF
Finished File transfer finished notification (receiver to sender). FIN
ACK EOF Positive acknowledgment of the EOF. ACK(EOF)
ACK Finished Positive acknowledgment of the Finished PDU. ACK(FIN)
NAK Negative acknowledgment by requesting retransmission(s) of missing data.

CFDP OVERVIEW

CFDP IN A NUTSHELL

- Class 1 (Open loop)
  - No retransmission, no ACK/NAK
  - No uplink required at ground stations
  - File lost if packet lost (& no on-ground reconstruction capabilities)



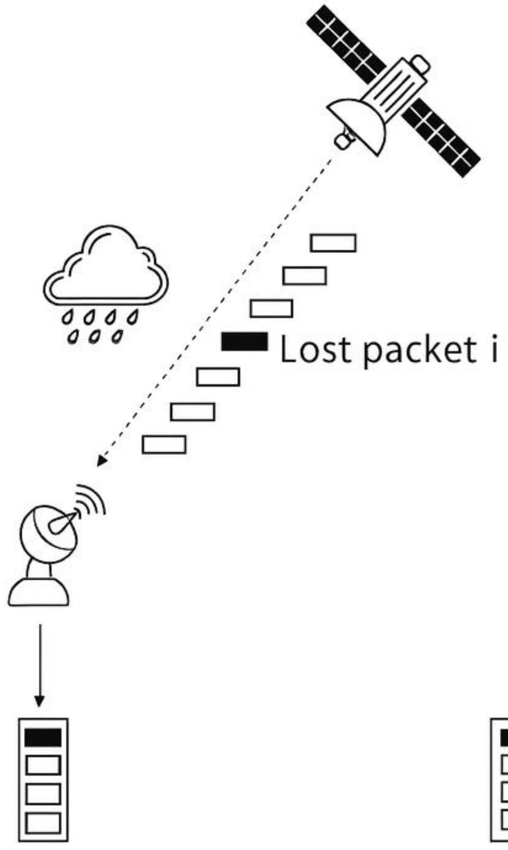
PDU -> Protocol Data Units
Metadata Attached data transferred with file data itself. M
File data A given portion (Nth) of the file data to transfer. FD(N)
EOF The End Of File notification (sender to receiver). EOF
Finished File transfer finished notification (receiver to sender). FIN

METHODOLOGY

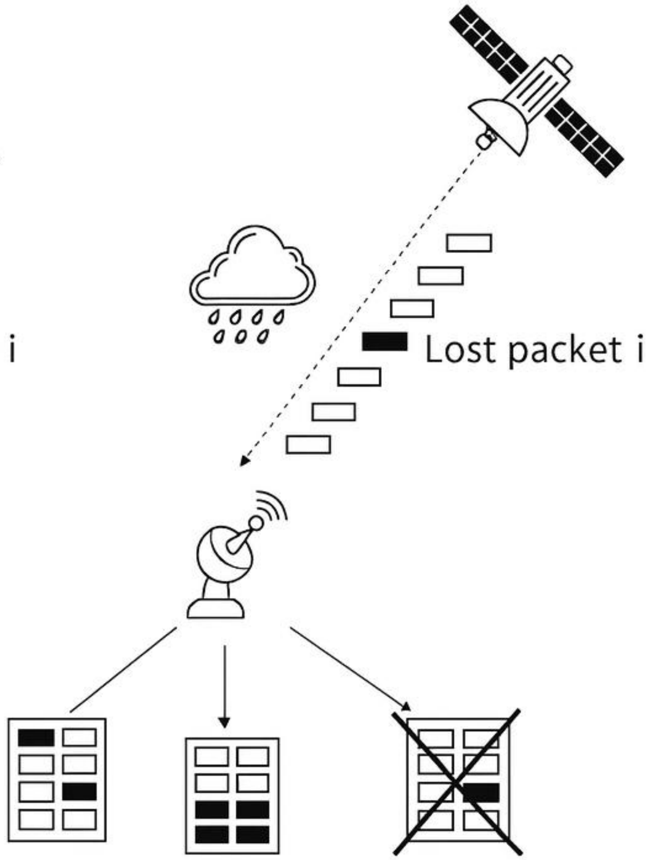
HOW WE APPROACH FILE SIZE OPTIMIZATION

- Introduction of „Equivalent Link Availability“

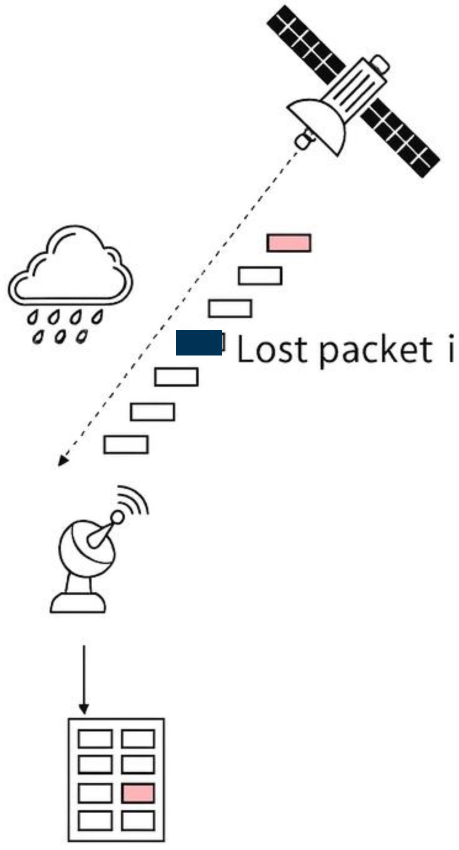
- Packet Store (no CFDP)



- Class 1



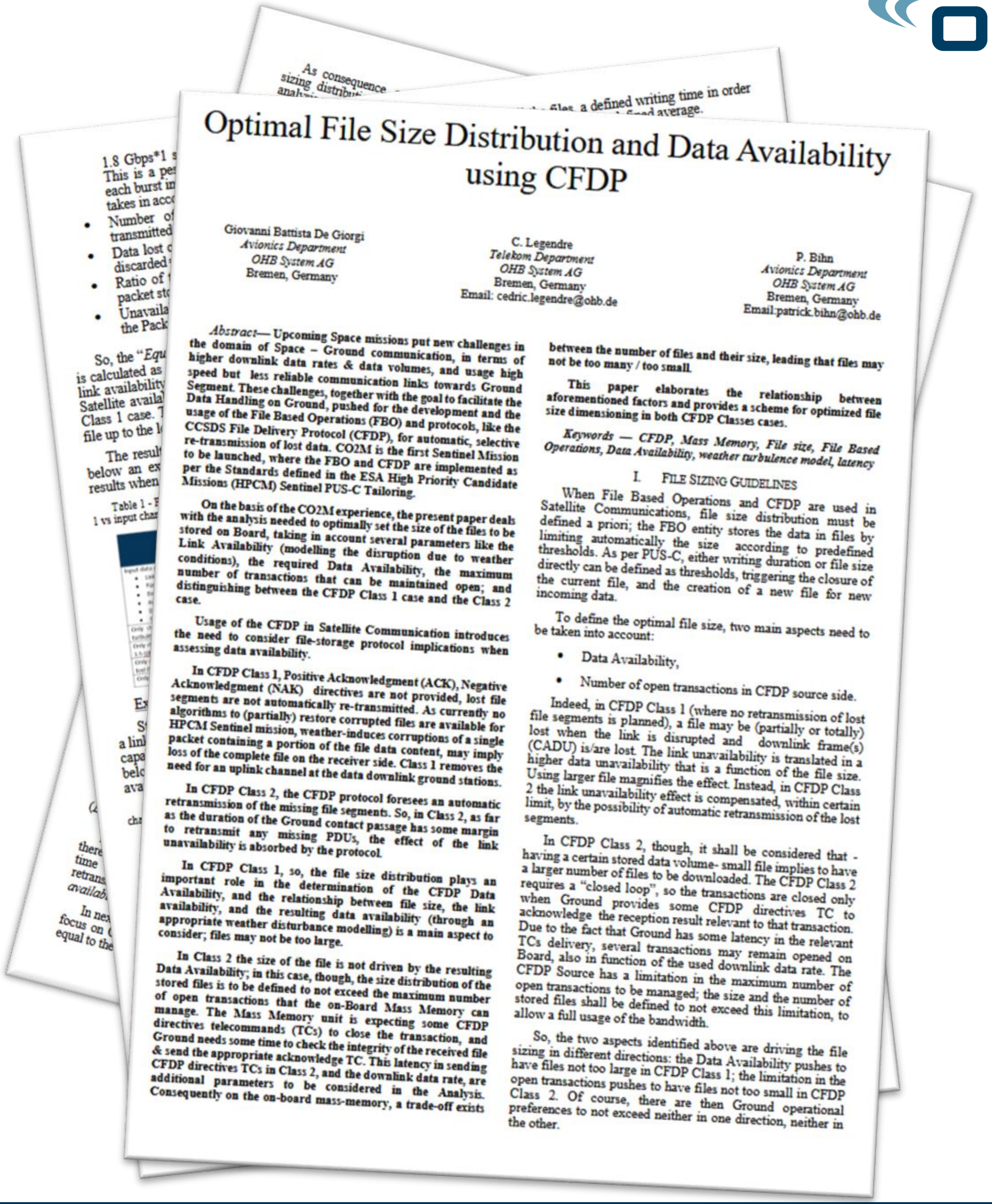
- Class 2



OBJECTIVES OF THIS TASK

WHAT THIS WORK ADDRESSES

- Define optimal file size distribution for missions using CFDP
- Consider key mission parameters:
  - Link availability
  - Required mission data availability
  - On-Board mass memory constraints
  - Ground latency for CFDP directives
- Compare Class 1 vs. Class 2 trade-offs
- Provide guidelines for configuring on-board file sizes



Optimal File Size Distribution and Data Availability using CFDP

Giovanni Battista De Giorgi
Avionics Department
OHB System AG
Bremen, Germany

C. Legendre
Telekom Department
OHB System AG
Bremen, Germany
Email: cedric.legendre@ohb.de

P. Bihn
Avionics Department
OHB System AG
Bremen, Germany
Email:patrick.bihn@ohb.de

Abstract— **Upcoming Space missions put new challenges in the domain of Space – Ground communication, in terms of higher downlink data rates & data volumes, and usage high speed but less reliable communication links towards Ground Segment. These challenges, together with the goal to facilitate the Data Handling on Ground, pushed for the development and the usage of the File Based Operations (FBO) and protocols, like the CCSDS File Delivery Protocol (CFDP), for automatic, selective re-transmission of lost data. CO2M is the first Sentinel Mission to be launched, where the FBO and CFDP are implemented as per the Standards defined in the ESA High Priority Candidate Missions (HPCM) Sentinel PUS-C Tailoring.**

On the basis of the CO2M experience, the present paper deals with the analysis needed to optimally set the size of the files to be stored on Board, taking in account several parameters like the Link Availability (modelling the disruption due to weather conditions), the required Data Availability, the maximum number of transactions that can be maintained open; and distinguishing between the CFDP Class 1 case and the Class 2 case.

Usage of the CFDP in Satellite Communication introduces the need to consider file-storage protocol implications when assessing data availability.

In CFDP Class 1, Positive Acknowledgment (ACK), Negative Acknowledgment (NAK) directives are not provided, lost file segments are not automatically re-transmitted. As currently no algorithms to (partially) restore corrupted files are available for HPCM Sentinel mission, weather-induces corruptions of a single packet containing a portion of the file data content, may imply loss of the complete file on the receiver side. Class 1 removes the need for an uplink channel at the data downlink ground stations.

In CFDP Class 2, the CFDP protocol foresees an automatic retransmission of the missing file segments. So, in Class 2, as far as the duration of the Ground contact passage has some margin to retransmit any missing PDUs, the effect of the link unavailability is absorbed by the protocol.

In CFDP Class 1, so, the file size distribution plays an important role in the determination of the CFDP Data Availability, and the relationship between file size, the link availability, and the resulting data availability (through an appropriate weather disturbance modelling) is a main aspect to consider; files may not be too large.

In Class 2 the size of the file is not driven by the resulting Data Availability; in this case, though, the size distribution of the stored files is to be defined to not exceed the maximum number of open transactions that the on-Board Mass Memory can manage. The Mass Memory unit is expecting some CFDP directives telecommands (TCs) to close the transaction, and Ground needs some time to check the integrity of the received file & send the appropriate acknowledge TC. This latency in sending CFDP directives TCs in Class 2, and the downlink data rate, are additional parameters to be considered in the Analysis. Consequently on the on-board mass-memory, a trade-off exists between the number of files and their size, leading that files may not be too many / too small.

This paper elaborates the relationship between aforementioned factors and provides a scheme for optimized file size dimensioning in both CFDP Classes cases.

Keywords — CFDP, Mass Memory, File size, File Based Operations, Data Availability, weather turbulence model, latency

I. FILE SIZING GUIDELINES

When File Based Operations and CFDP are used in Satellite Communications, file size distribution must be defined a priori; the FBO entity stores the data in files by limiting automatically the size according to predefined thresholds. As per PUS-C, either writing duration or file size directly can be defined as thresholds, triggering the closure of the current file, and the creation of a new file for new incoming data.

To define the optimal file size, two main aspects need to be taken into account:

- Data Availability,
- Number of open transactions in CFDP source side.

Indeed, in CFDP Class 1 (where no retransmission of lost file segments is planned), a file may be (partially or totally) lost when the link is disrupted and downlink frame(s) (CADU) is/are lost. The link unavailability is translated in a higher data unavailability that is a function of the file size. Using larger file magnifies the effect. Instead, in CFDP Class 2 the link unavailability effect is compensated, within certain limit, by the possibility of automatic retransmission of the lost segments.

In CFDP Class 2, though, it shall be considered that - having a certain stored data volume- small file implies to have a larger number of files to be downloaded. The CFDP Class 2 requires a “closed loop”, so the transactions are closed only when Ground provides some CFDP directives TC to acknowledge the reception result relevant to that transaction. Due to the fact that Ground has some latency in the relevant TCs delivery, several transactions may remain opened on Board, also in function of the used downlink data rate. The CFDP Source has a limitation in the maximum number of open transactions to be managed; the size and the number of stored files shall be defined to not exceed this limitation, to allow a full usage of the bandwidth.

So, the two aspects identified above are driving the file sizing in different directions: the Data Availability pushes to have files not too large in CFDP Class 1; the limitation in the open transactions pushes to have files not too small in CFDP Class 2. Of course, there are then Ground operational preferences to not exceed neither in one direction, neither in the other.

RESULTS FOR CFDP CLASS 1

EXAMPLE (CLASS 1)

- Input
  - Link Availability: 99 %
  - Passage duration : 6 min
  - Downlink data rate: 1.8 Gbps
  - Average file size: 1 Gbyte
  - Duration of average turbulence:
    - 1 sec
  - Groundstation can reconstruct:
    - file up to the lost packet
  - Flat distribution

RESULTS FOR CFDP CLASS 1

EXAMPLE (CLASS 1)

- Equivalent Link Availability:

1. Time of missing communication: $T_{loss} = \frac{(100-L_a)}{100} * \left(T_{pass} * 60\right) = \frac{1}{100} * (6 * 60) = 3{,}6\ seconds$
2. Data Lost in Packet store case: $D_{loss\_ps} = R_{dl} * T_{loss} = 1.8\ Gbps\ * 3.6\ \text{sec} = 6.5\ Gbit$ (0.8 GB)
3. Disruption / Burst events: $N_{event} = \left\lceil \frac{T_{loss}}{T_{Turb}} \right\rceil = \left[\frac{3{,}6\ sec}{1\ sec}\right] = 3.6 = \sim 4$
4. Number of Files transmitted per burst event $F_{event} = \left\lfloor \frac{R_{dl} * T_{turb}}{S_{file} * 8} + 1 \right\rfloor = \left\lfloor \frac{1.8 * 1}{1 * 8} + 1 \right\rfloor = 1{,}225 = \sim 2$
5. Number of impacted files during passage $F_{imp} = F_{event} * N_{event} = 2 * 4 = 8$
6. Data lost in Class 1: $D_{loss\_c1} = \frac{F_{imp} * S_{file}}{2} = \frac{8 * 1\ Gbyte}{2} = 4\ Gbyte$
7. Ratio of data lost $R_{Ratio} = \frac{D_{loss\_c1}}{D_{loss\_ps}} = \frac{4\ Gbit}{6.5\ Gbit} = 4.9$
8. Unavailability in Class1 $U_{C1} = R_{ratio} * \frac{(100-L_a)}{100} = 4{,}9 * \frac{(100-99)}{100} = 4{,}9\%$
9. Equivalent link availability CFDP Class 1 $ELA_{C1} = 1 - U_{C1} = 1 - 4{,}9\% = \mathbf{95{,}1\%}$
10. **Data Availability:** $DA_{Mission} = Satellite\ Availability * ELA_{C1}$

- Input
  - Link Availability: 99 %
  - Passage duration : 6 min
  - Downlink data rate: 1.8 Gbps
  - Average file size: 1 Gbyte
  - Duration of average turbulence:
    - 1 sec
  - Groundstation can reconstruct:
    - file up to the lost packet
  - Flat distribution

Symbol	Description	Unit
LU	Link unavailability	%
T_{pass}	Passage duration	s
T_{loss}	Time of missing communication during a passage	s
R_{dl}	Downlink data rate	Gbps
$D_{loss,PS}$	Data lost in packet store case	Gbit
T_{turb}	Average duration of a disruption event	s
N_{event}	Number of disruption events during a passage	–
S_{file}	Average file size	GB
F_{event}	Number of files transmitted per disruption event	–
F_{imp}	Total number of impacted files	–
$D_{loss,C1}$	Data lost in CFDP Class 1 case	GB
R_{ratio}	Ratio of data lost (Class 1 vs. packet store)	–
U_{PS}	Data unavailability in packet store case	%
U_{C1}	Data unavailability in CFDP Class 1	%

RESULTS FOR CFDP CLASS 1

EXAMPLE (CLASS 1)

- Further calculations:

  - **Conclusion: The average file size in Class 1 shall be chosen in order that the resulting „ ELA_{C1} * Satellite Availabilty“, is compliant to the Mission data availability requirement.**

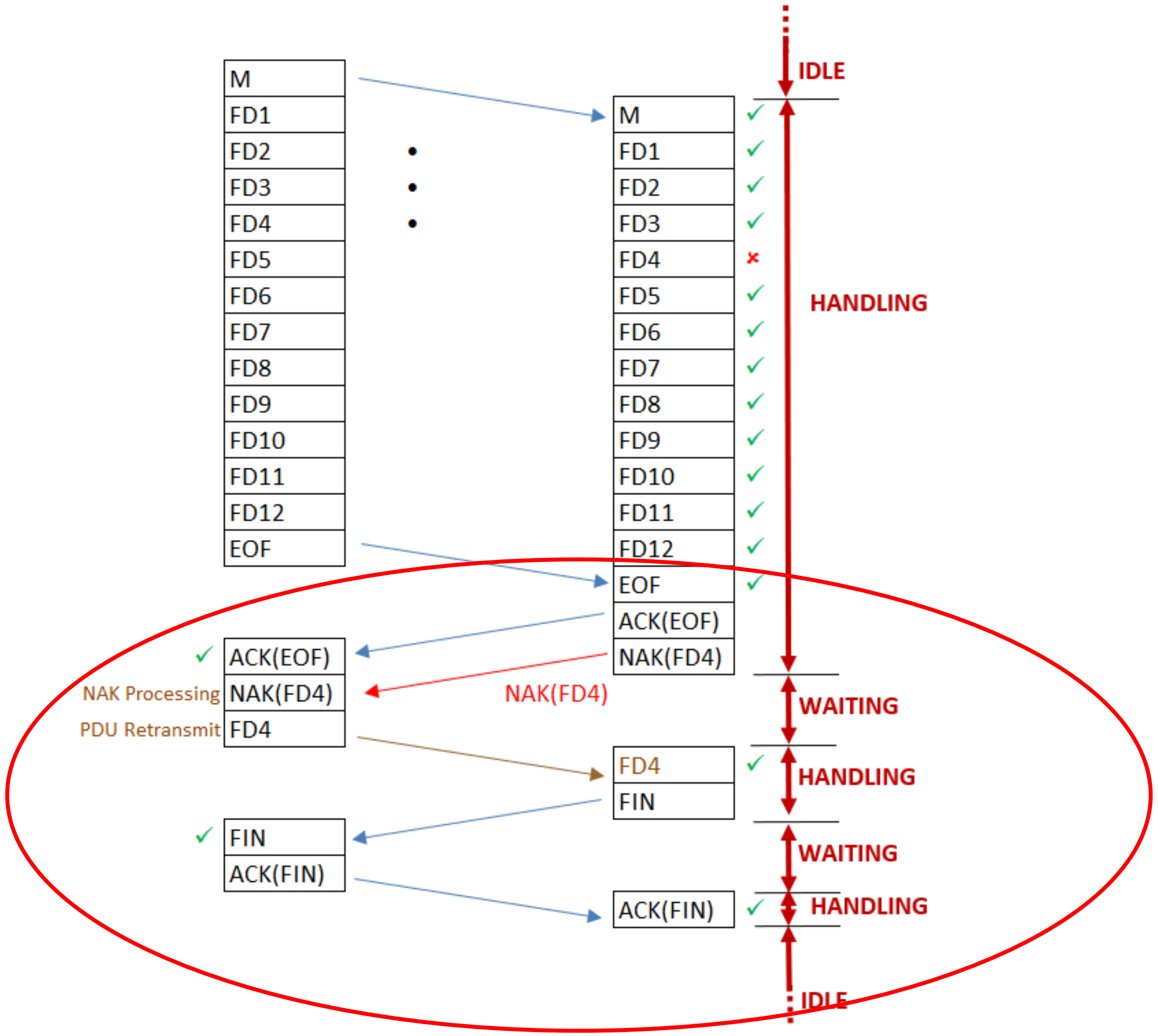
- Can be reached by either:
  – a file size equal to the defined average
  – a defined writing time in order to have an average size equal to the defined average

Input Data	Equivalent link availability CFDP Class1
Input data and assumption in case 1: • Link availability: 99.8% • Pass duration: 6 min • Downlink data rate: 1.8 Gbps • Average file size: 1 GByte • Duration of average weather turbulence: 1 sec • Ground is able to reconstruct part of the file	95.1%
Only changed input with respect to case 1: “Duration of average weather turbulence”, from 1 to 3 sec	97.5%
Only changed input with respect to case 1: “Average File size”, from 1 GByte to 1.5 Gbyte	92.6%
Only changed input with respect to case 1: Ground discards the whole file in case of a lost PDU	90.1%
Only changed input with respect to case 1: Ground reconstructs the whole file	99%

RESULTS FOR CFDP CLASS 2

TRANSACTION TIMING MODEL

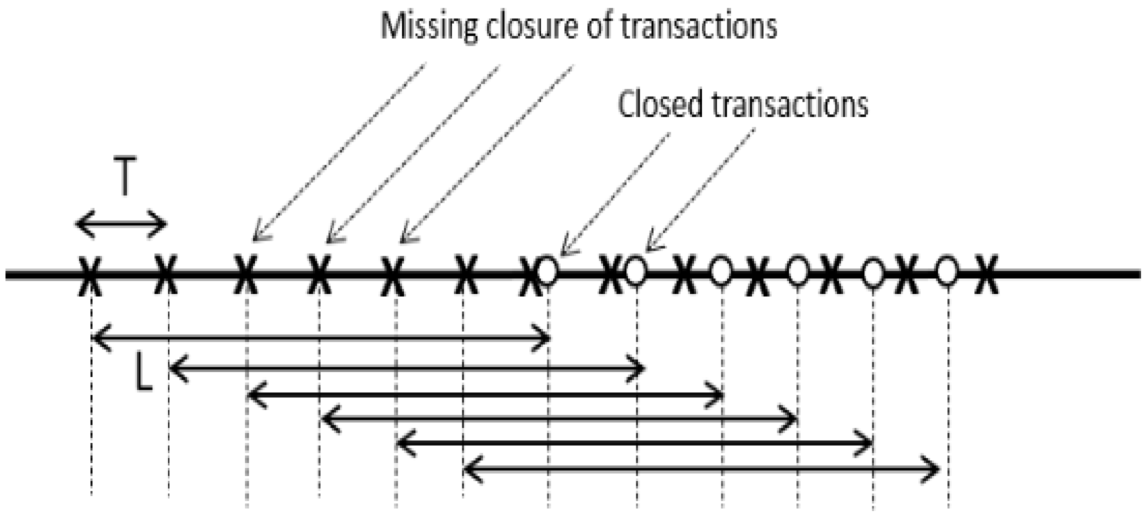
- Class 2 (closed loop):
  - Equivalent link availability of 1
    - IF there is a positive link budget and time margin
- Smaller Files create more overhead
  - Higher Number of transactions
  - Open transactions limited by CFDP Source (PDHU, OBC…)
  - Latency of closure due to Ground TCs
- Better to use large files (?)



RESULTS FOR CFDP CLASS 2

EXAMPLE (CLASS 2)

- Assumptions for Class 2
  - File sizes are averaged
  - Interleaved
    - Better use of the downlink
    - Smoother transaction management
    - Reduced risk of reaching open transaction limits
- Latency:
  - total round-trip delay
    - onboard queuing and transmission delay
    - propagation, reception, and processing delay on the ground
    - the uplink delay for the acknowledgment to reach the spacecraft
- Goal:
  - **Optimize for a medium number of not closed transactions.**



RESULTS FOR CFDP CLASS 2

EXAMPLE (CLASS 2)

1. Duration of downlink time in minutes $P_{dl} = \frac{\left[\frac{V_{data}}{R_{dl}}\right]}{60}$
2. Time to close transcations $Tr_{cl} = P_{dl} * \frac{60}{N_{files}}$
3. Not closed transactions $Tr_{nc} = (L/T_{cl}) + 1$
   – Worst Case (peak) $Tr_{nc\_p} = \left(L * N_{files} * \frac{R_{dlo}}{V_{data}}\right) + 1 + M_{files}$

V_{data} = Stored volume to be downloaded
R_{dl} = Data Download Speed [Gbit]

N_{files} = Number of files to be downlinked in one orbit (Mission data)
L = Latency [seconds]

M_{files} = Number of files to be downlinked (out of family, e.g. TM from OBC)

- N_{files} = 100; M_{files}= 10; V_{data} = 534 Gb; R_{dl} = 1.8 Gbps

Total Delay	Tr_{nc}	Tr_{nc_p}
3 sec	3	13
6 sec	4	14
8 sec	4	14
13 sec	6	16

CONCLUSIONS & OPEN WORK

- Conclusions
  - Class 1:
    - The Equivalent Link Availability quantifies how file size and link outages affect effective data availability.
    - Smaller files reduce the amplification of physical losses.
  - Class 2:
    - Data availability is near 100 %; performance depends on latency (Tr_{cl}) and the number of open transactions.
    - A small, stable number of open transactions confirms proper CFDP operation.
  - Together, these models provide a comprehensive framework for analyzing and comparing CFDP performance.
  - File size depends on mission needs, Bandwidth and also Instrument / Onboard Data rates.
- Open Work
  - Extend the model to variable file size distributions (non-uniform file sets).
  - Include dynamic channel conditions (time-varying turbulence statistics).
  - Implement real CFDP simulation to validate analytical assumptions.
  - Investigate automatic file segmentation/adaptation based on real-time link feedback.

ANY FURTHER POINTS OR IDEAS?

CONTACT US AT: **HTTPS://WWW.OHB.DE/EN/CAREER**

or Patrick.Bihn@me.com

G. B. DE GIORGI, C. LEGENDRE, P. BIHN

PATRICK BIHN, 15.10.2025

THANK YOU!

OHB SE
Manfred-Fuchs-Platz 2-4
28359 Bremen
Germany