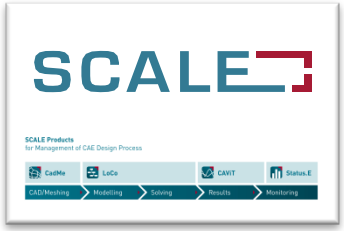
NEW DEVELOPMENTS ON COMPRESSION AND TRANSFER OF SIMULATION DATA WITHIN AN SDM SYSTEM

Matthias Büchse, M. Thiele, H. Müllerschön
SCALE GmbH, Germany

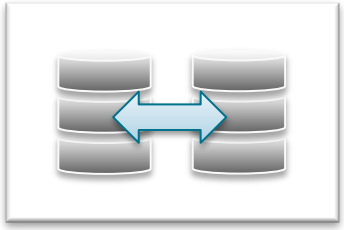
SCALE

Agenda



Company and Products - Brief Introduction

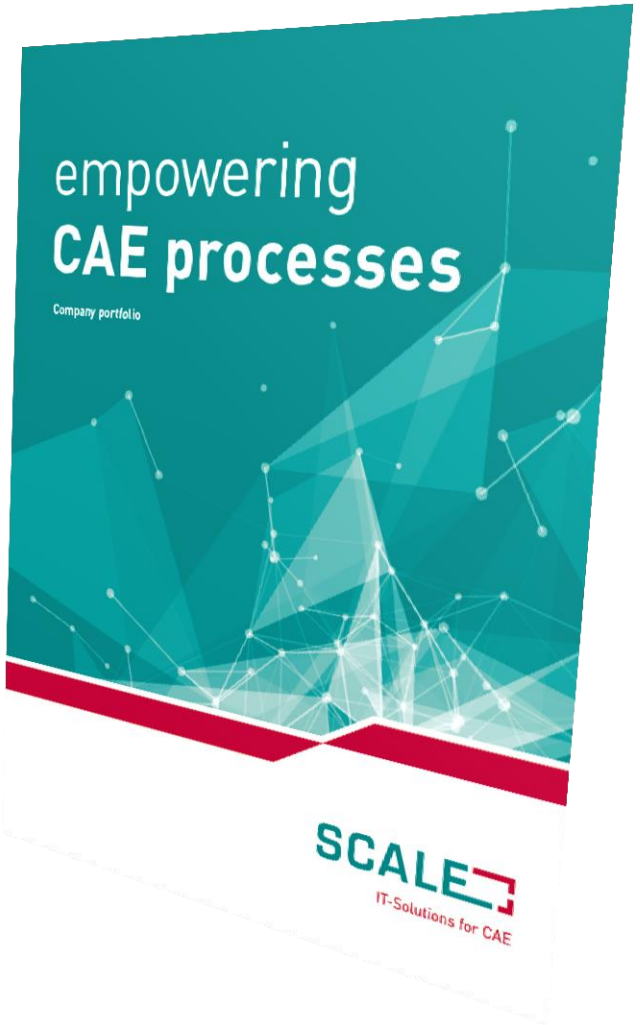
Motivation for Data Compression



Compression using Data Deduplication

SCALE GmbH

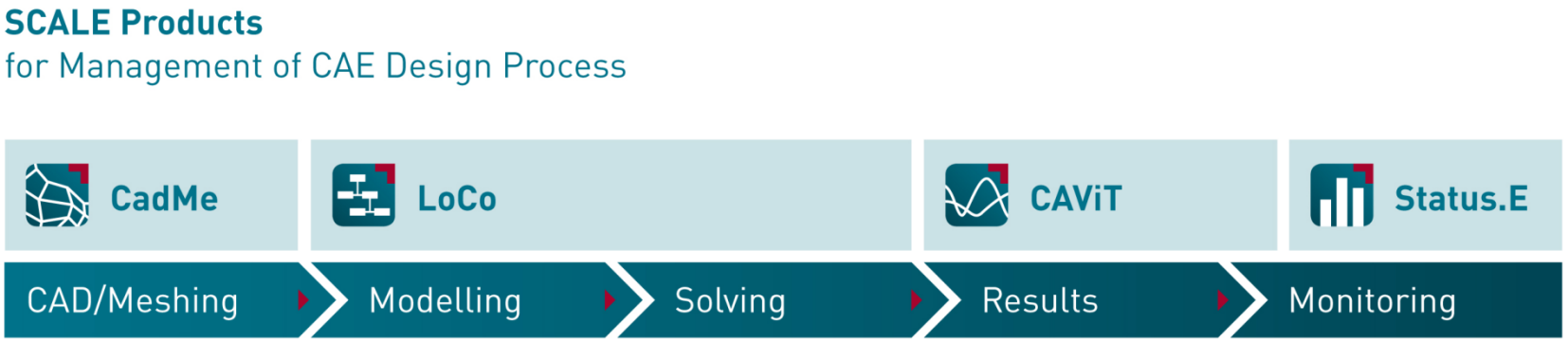
- Company is dedicated to „CAE process-, and data management“
- SCALE is a100% subsidiary of DYNAmore
- Currently ~35 people
  (CAE-engineers and computer scientists)
- Offices in Germany
  - Ingolstadt
  - Stuttgart
  - Wolfsburg
  - Dresden (Software development)
- International partners in cooperation with DYNAmore Group



SCALE

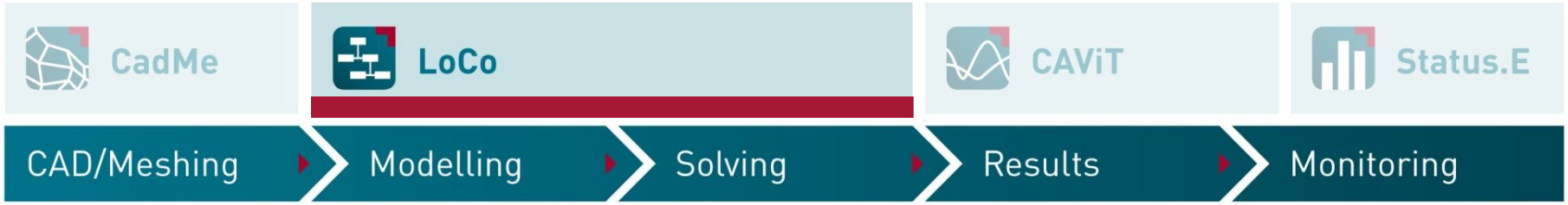
SCALE Products

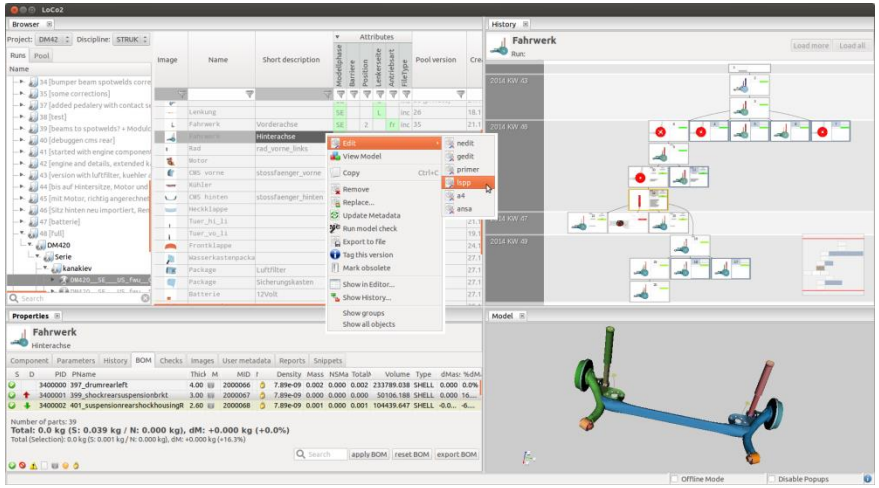
- SCALE has developed a comprehensive simulation and test data framework (SCALE.sdm) in close collaboration with Volkswagen Group (AUDI, Porsche, Volkswagen, Seat).
- Several Apps cover the entire CAE design process

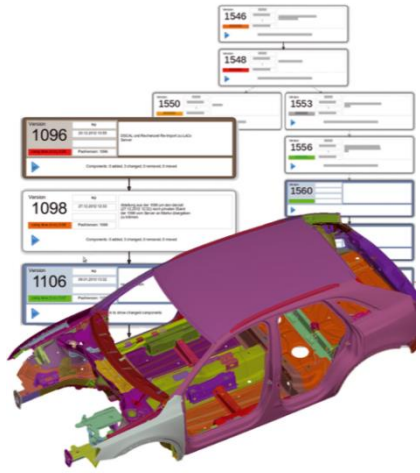


- The system is running today with more than 800 registered users at VW Group
- This presentation focuses on LoCo: SCALE's software application for simulation model data management

SDM - Application LoCo







- Simulation Data- / Variant Management
  - Workbench for Simulation Engineers
  - Unique RichClient/Offline-concept with sync-mechanism *(internal/external)*
- Workflows / Features
  - Integration of any third party or in-house CAE-product
  - Solver: PAM-Crash, LS-DYNA, Nastran, Abaqus, …
  - Job submit and monitoring
  - Optimization, robustness, DOE, …
  - Quality checks of models
  - Advanced security features
    - Two factor authentication
    - Encryption
  - Distributed, collaborative work environment
  - Access-, roles and rights management
  - Version Control

SCALE

Agenda

Company and Products - Brief Introduction

Motivation for Data Compression

Compression using Data Deduplication

Motivation: *SDM Data Dimensions at VW Group*

- Requirements
  - Legislation
  - Consumer tests
  - Customer comfort requirements

- Projects and derivatives
  - Body variants
  - Engine variants
  - Interior configuration
  - Region specifics

- Collaborative development
  - VW Group – many brands
  - Engineering Service partners
  - Suppliers

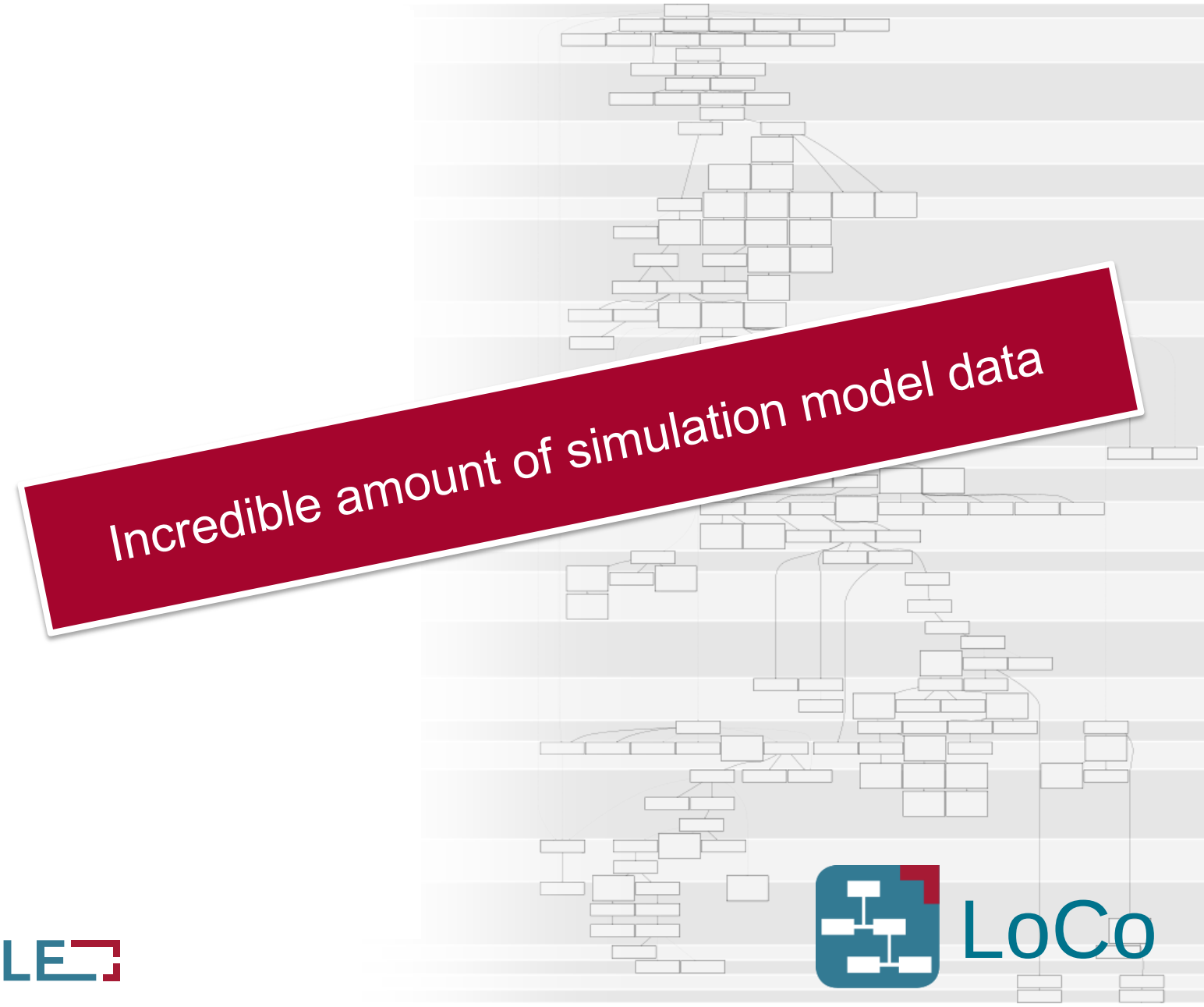
Environment diversity

Product diversity

Collaboration diversity

SCALE

Motivation: Growing amounts of data/simulations



Incredible amount of simulation model data

SCALE

LoCo

Motivation: *Location Diversity*

- Collaboration
  - Teams are distributed all over the world
  - Products share data over multiple sites
  - Many engineers are working together on the same problem
- Availability
  - Users expect data to be instantly available
  - Bandwidth and latency are critical
- Security
  - Encryption is essential

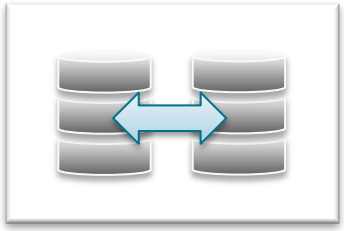
- external partners
- sites

SCALE

Agenda

Company and Products - Brief Introduction

Motivation for Data Compression

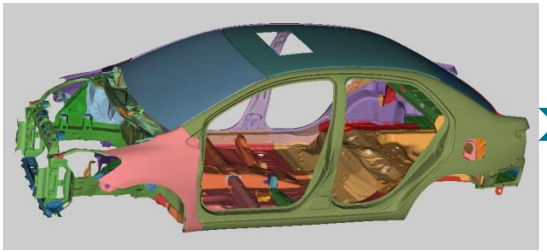


Compression using Data Deduplication

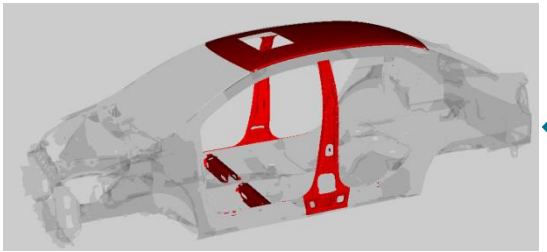
Simulation Data Management Workflow - TODAY

Server

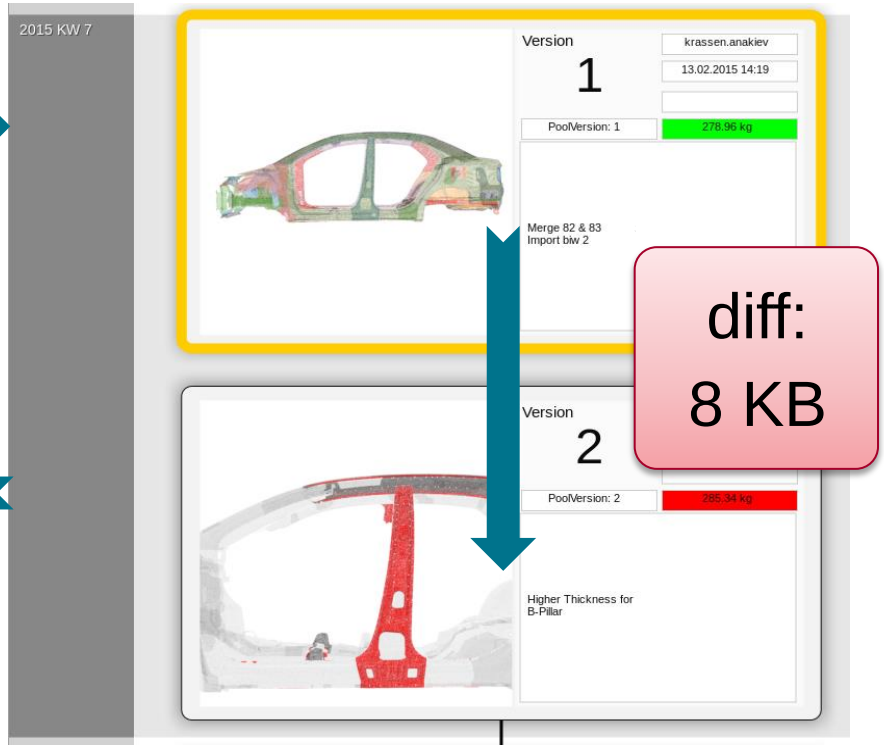
Client (local)



140 MB



140 MB



diff:
8 KB

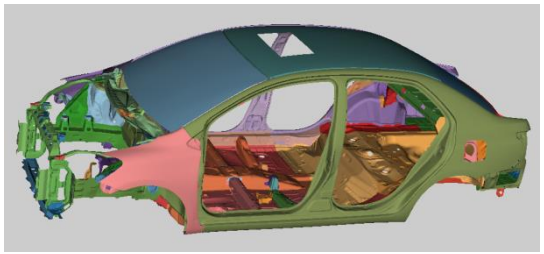
storage 280 MB

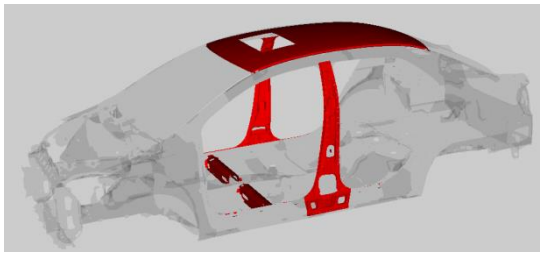
storage 280 MB

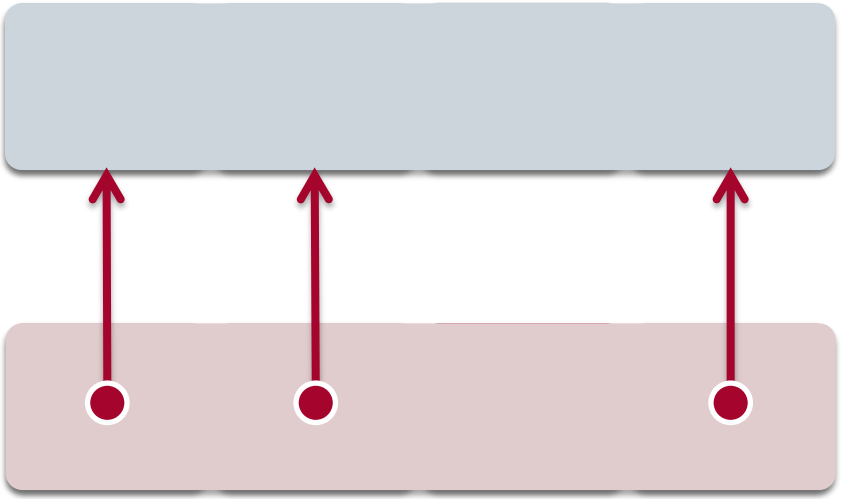
transfer 280 MB

SCALE

Data Deduplication: Approach







Chunking: find block boundaries via rolling checksum

Indexing: identify each block with cryptographic hash

Data Deduplication: Approach

- initial file

`LoCo_speichert_nur_das_was_nötig_ist.`

Block A: `LoCo_spe`
Block B: `ichert_n`
Block C: `ur_das_w`
Block D: `as_nötig`
Block E: `_ist.`

File consists of blocks: A B C D E

5 + 37 = 42 characters

- changed file

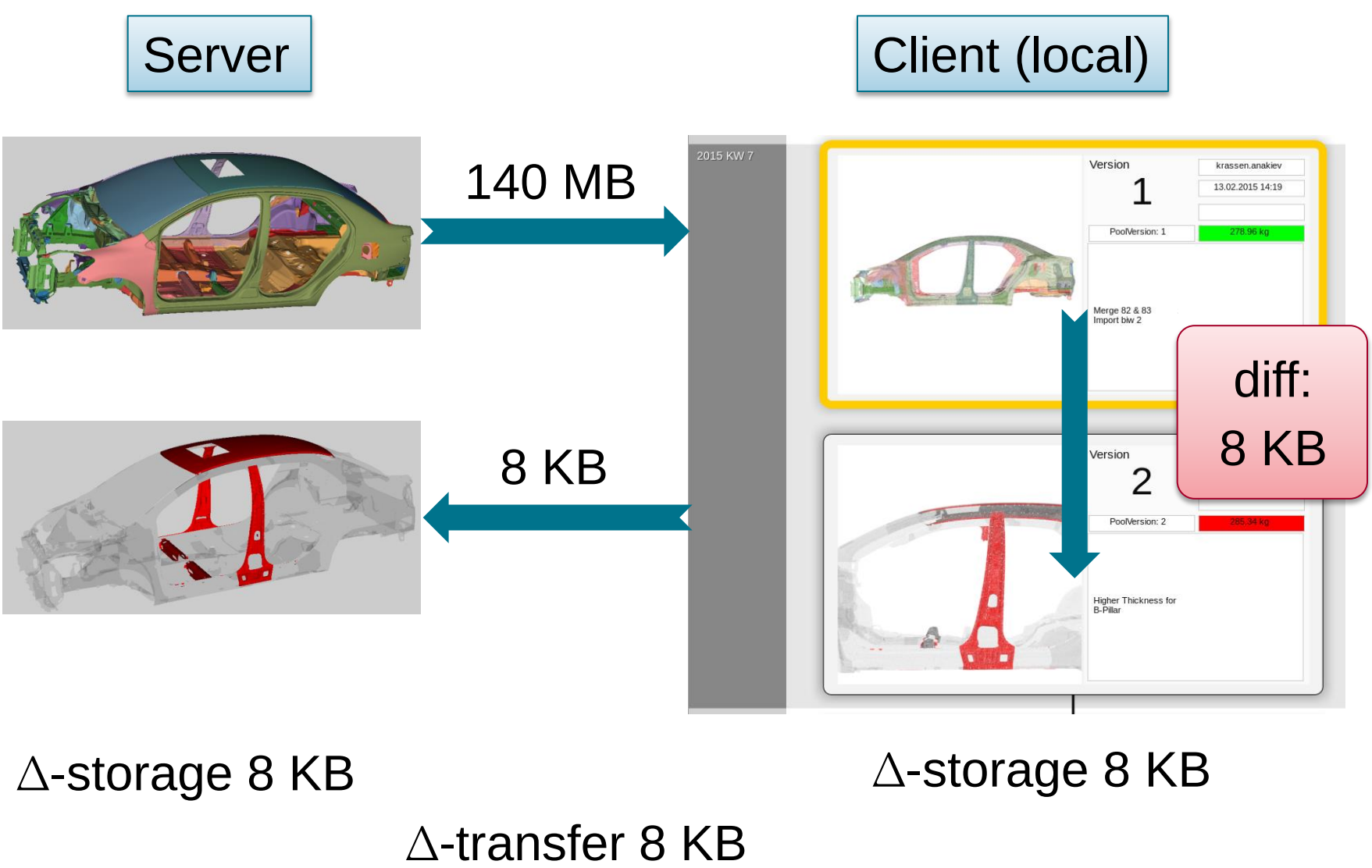
`LoCo_speichert_nur_das_was_geändert_ist.`

Block F: `as_geänd`
Block G: `ert`

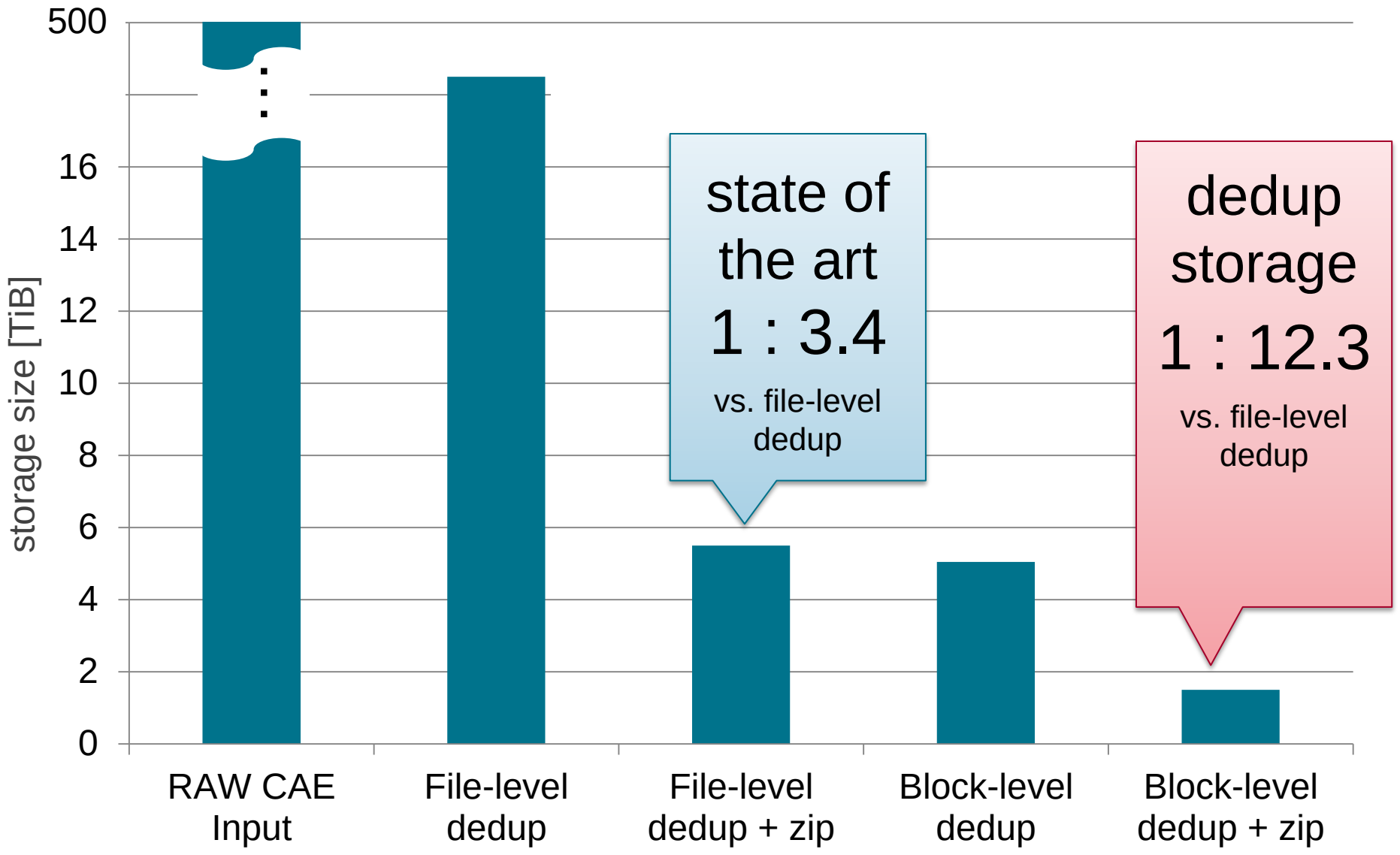
File consists of blocks: A B C E F G

6 + 11 = 17 characters

Simulation Data Management Workflow - TOMORROW

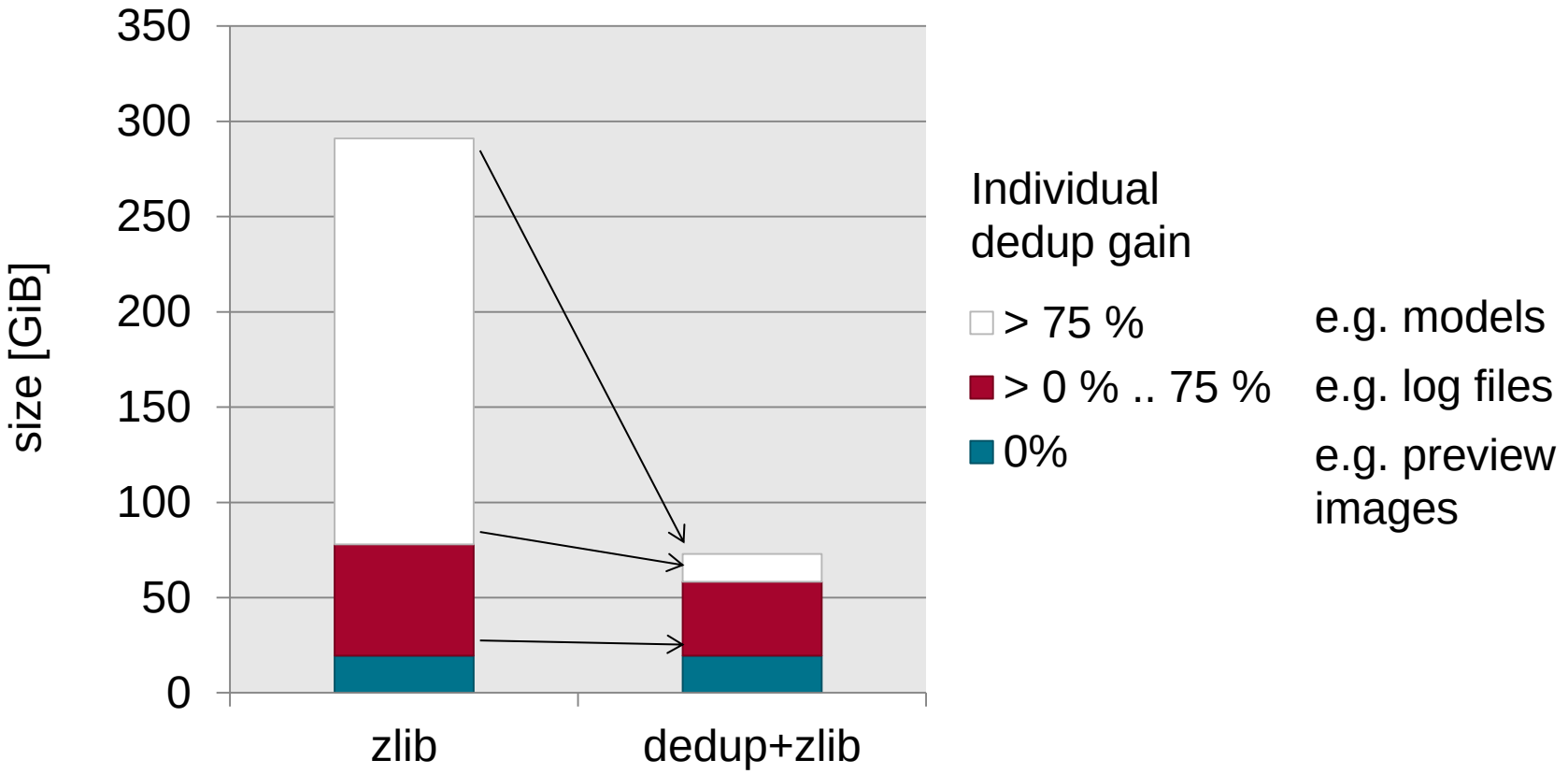


Data Deduplication: Real-World Car Project Data



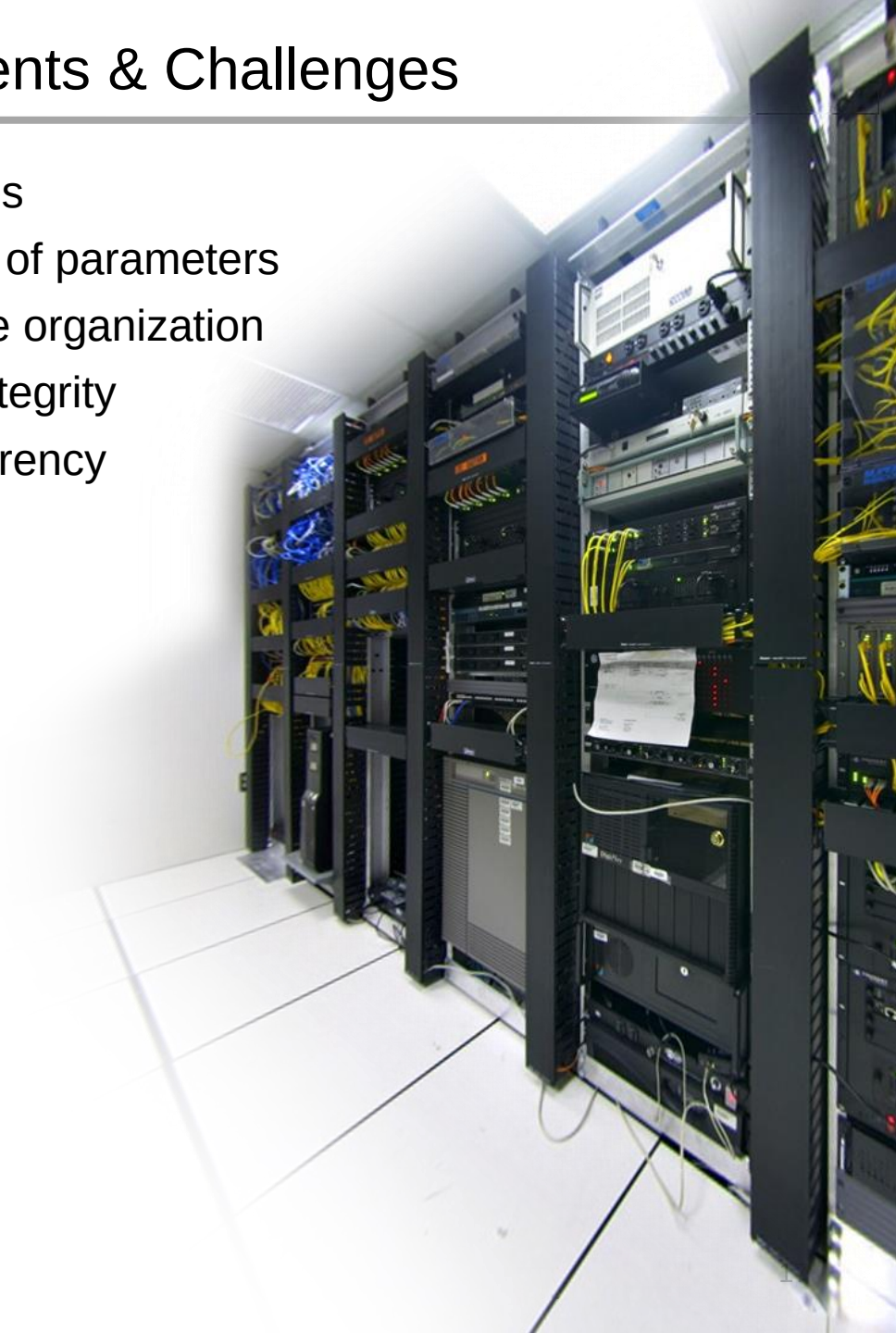
Data Deduplication: Results

- Example Data Vault
  - 280 GiB real-world zlib compressed data
  - Total deduplication ratio: 1 : 4



Data Deduplication: Requirements & Challenges

- Requirements
  - Minimized Storage
  - Minimized Transfer
  - Performance
  - Scalability
  - Deletion
  - Encryption
- Challenges
  - Choice of parameters
  - Storage organization
  - Data integrity
  - Concurrency



SCALE

Conclusions / Roadmap

- Done
  - Implementation of data deduplication in SCALE's SDM client *LoCo*
  - Significantly reduced storage of redundant data
    - savings compared to raw model data: 99,7%
    - savings compared to previous state of the art: 75 %
  - Encryption of data possible (similar to OpenPGP)

- Work in progress
  - Implementation of new technology into SCALE.SDM server (2017)
  - Significant reduction of transfer volume: Only transfer of deduplicated data (2018)

- Acknowledgements
  - The work on data deduplication has been developed in the big data project VAVID, which is funded by the German ministry of education and research