Automated Data Management
of SLR Data and Products
at the EUROLAS Data Centre (EDC)

Christian Schwatke

17th Workshop on Laser Ranging
Bad Kötzting, Germany, May 16-20 2011

International Laser Ranging Service

DGFI

Overview

- Past, present, future developments at the EDC

- System architecture

- Data flow within the ILRS and at the EDC

- ILRS - Mailing lists

- EDC - Website

C.Schwatke, Automated Data Management of SLR Data and Products at the EDC

Past, Present, Future

Timeline: 1994 ... 2009 2010 2011 2012

Server (old)
Server I (new)
Server II (new)
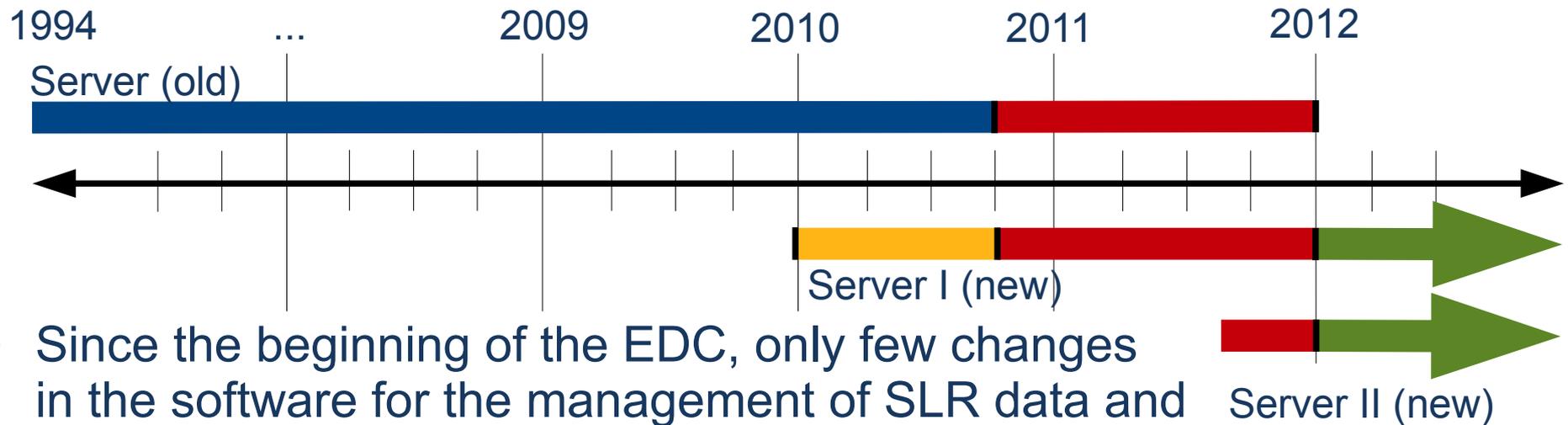
▶ Since the beginning of the EDC, only few changes in the software for the management of SLR data and products were made

▶ An upgrade from the old software to a new hardware respectively new operating system is hard to manage

▶ The old system was operated semi-automatic

▶ Increasing amount of data

▶ Requirement of hourly data exchange with CDDIS/ITT

Objective: Develop of a new software for the management of SLR-data and products

C.Schwatke, Automated Data Management of SLR Data and Products at the EDC

Past, Present, Future

1994 ... 2009 2010 2011 2012

Server (old)

Server I (new)

Server II (new)
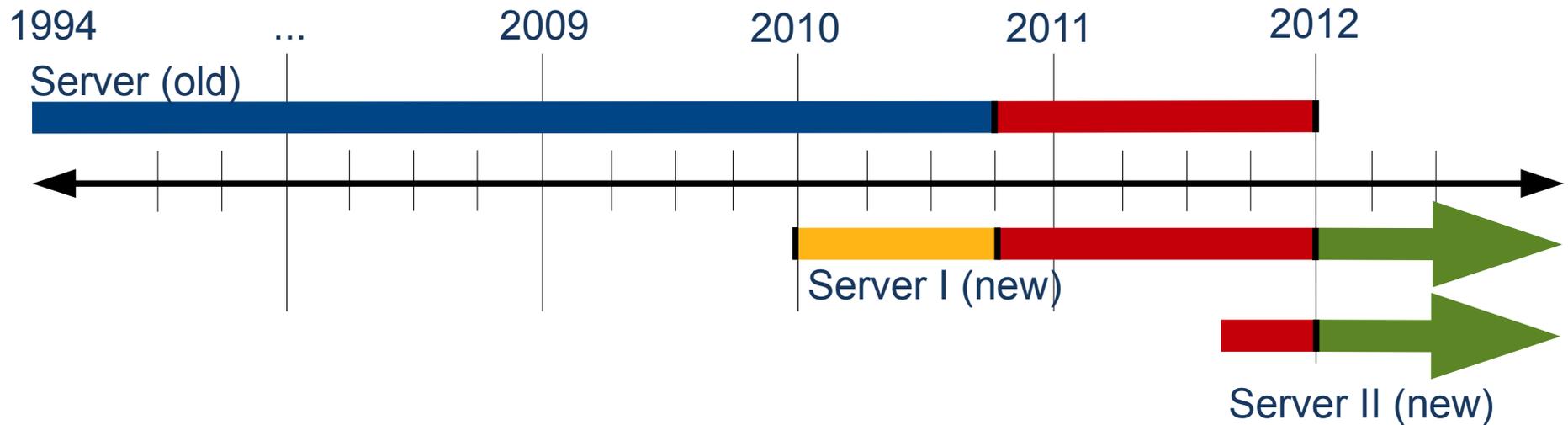
- ► Software development started at the beginning of 2010

- ► Changeover to the new system in Fall 2010

- ► Further software development has to be done
  with a view to the construction of a high-availibility cluster

- ► At the beginning of 2012

  - ► Shut down of the old server

  - ► Software development completed

C.Schwatke, Automated Data Management of SLR Data and Products at the EDC

System Architecture (past)
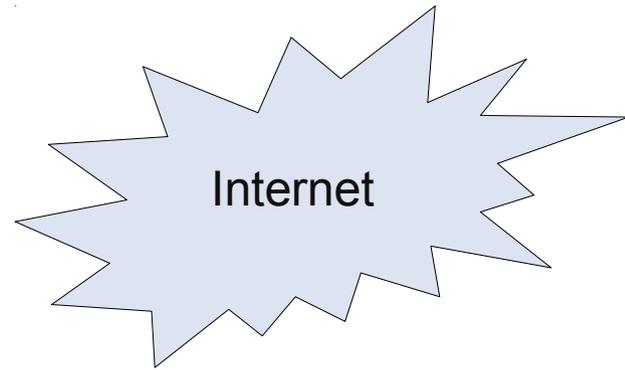
Internet

dgfi3.dgfi.badw-muenchen.de
129.187.165.3

- Manual backup on internal server.
- No backup server available for the operational system

C.Schwatke, Automated Data Management of SLR Data and Products at the EDC

System Architecture (present, future)

Internet

edc.dgfi.badw.de

- Minimisation of down time through port forwarding and changing IP addresses

- Backup for the operational system and data

edc1.dgfi.badw.de
129.187.165.11

edc2.dgfi.badw.de
129.187.165.12

Internal synchronisation (data, WWW, database)

C.Schwatke, Automated Data Management of SLR Data and Products at the EDC
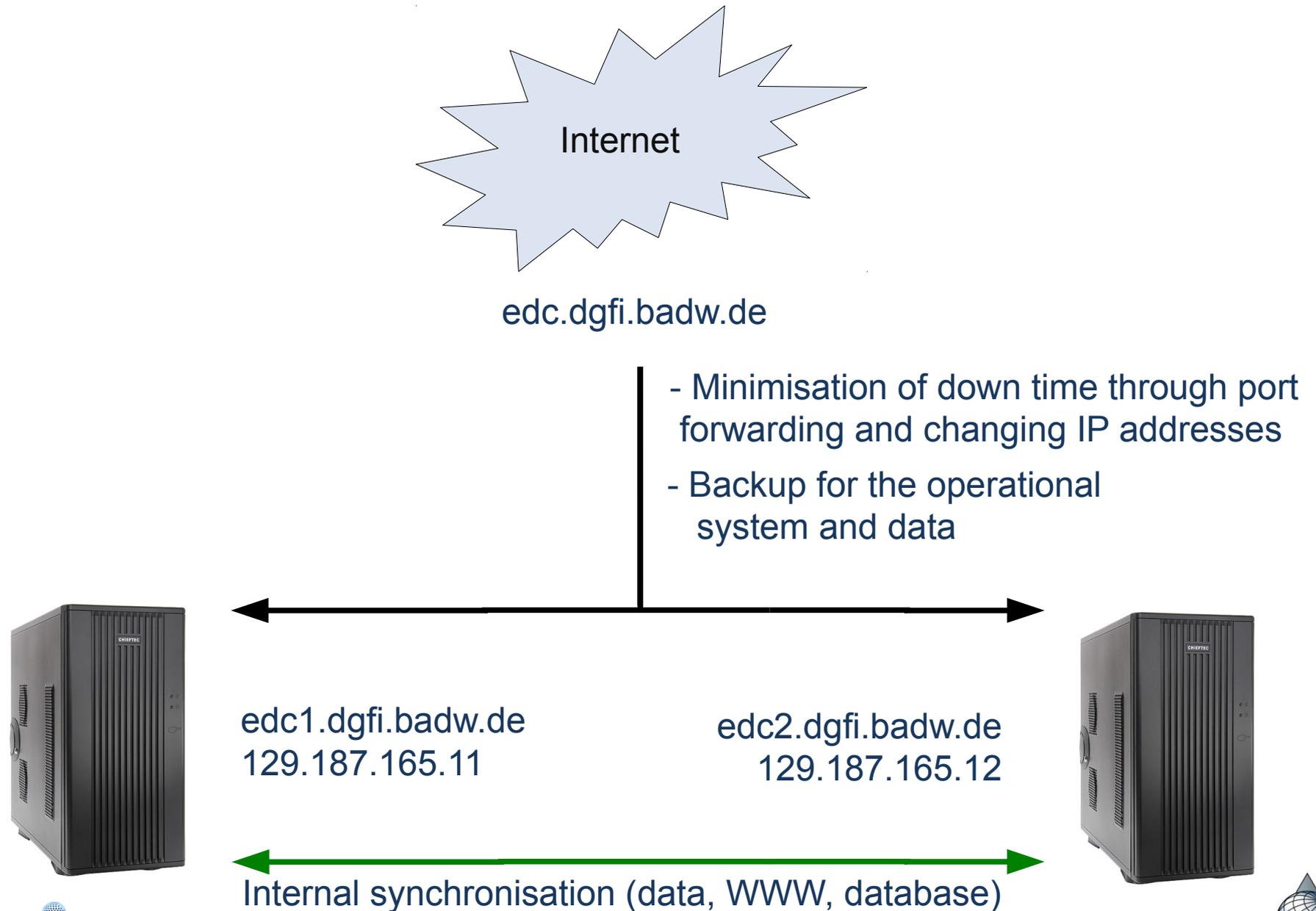
System Architecture (present, future)

Very Important!

EDC-Access:
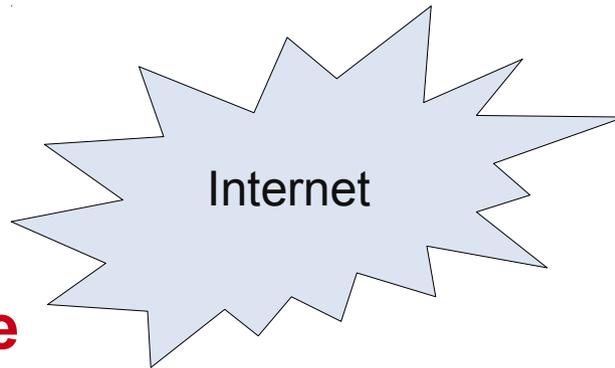
http://edc.dgfi.badw.de

Don't use !!!:
http://129.187.165.3
http://129.187.165.11
http://129.187.165.12
http://dgfi3.dgfi.badw-muenchen.de
http://edc1.dgfi.badw.de
http://edc2.dgfi.badw.de

Internet

edc.dgfi.badw.de

- Minimisation of down time through port forwarding and changing IP addresses

- Backup for the operational system and data

edc1.dgfi.badw.de
129.187.165.11

edc2.dgfi.badw.de
129.187.165.12

Internal synchronisation (data, WWW, database)

C.Schwatke, Automated Data Management of SLR Data and Products at the EDC

International Laser Ranging Service

DGFI

Data Flow within the ILRS

C.Schwatke, Data Formats & Procedures Working Group

Data Flow at EDC

FTP
(Incoming)

Fetch Data

Mailbox

Incoming
Data
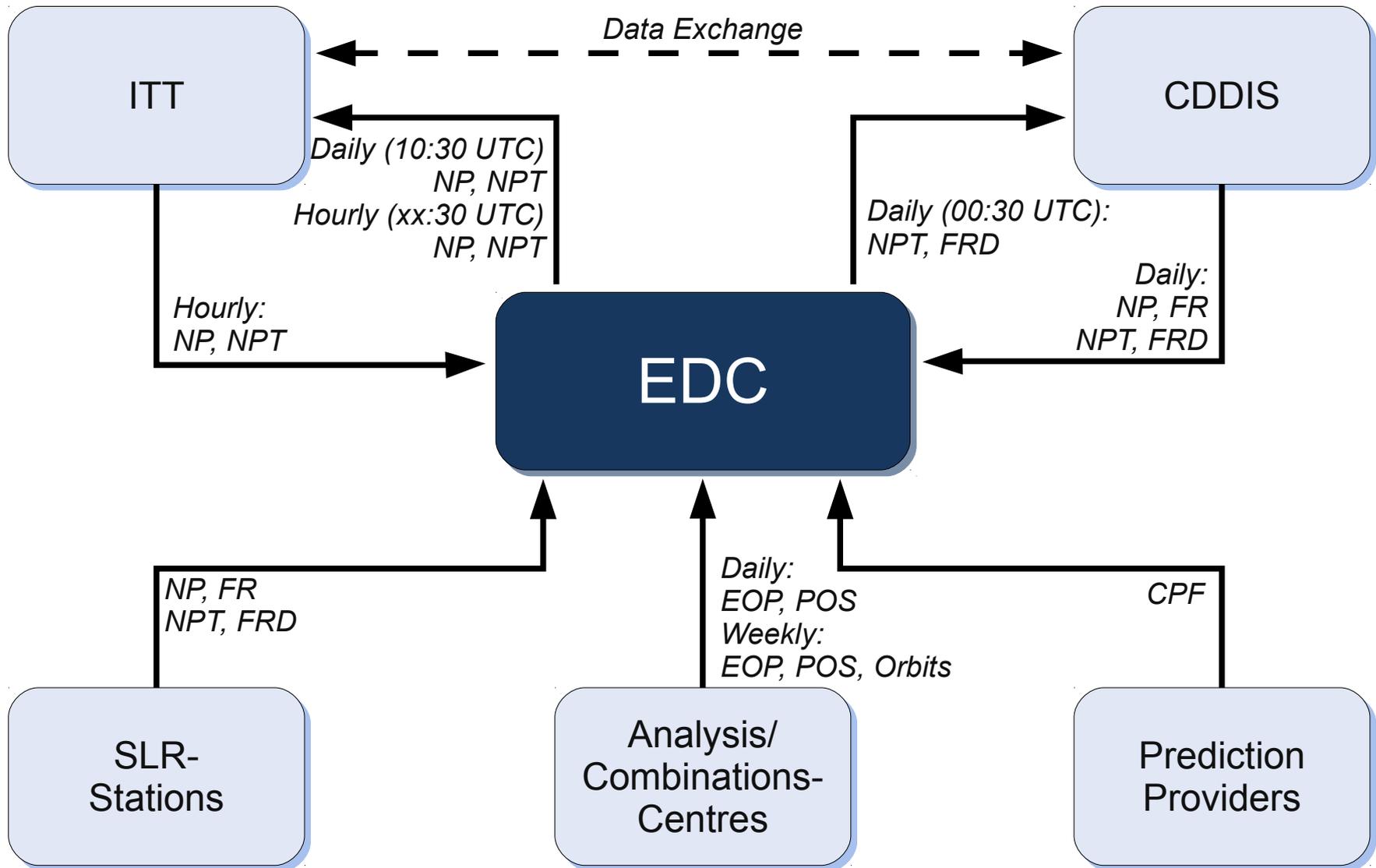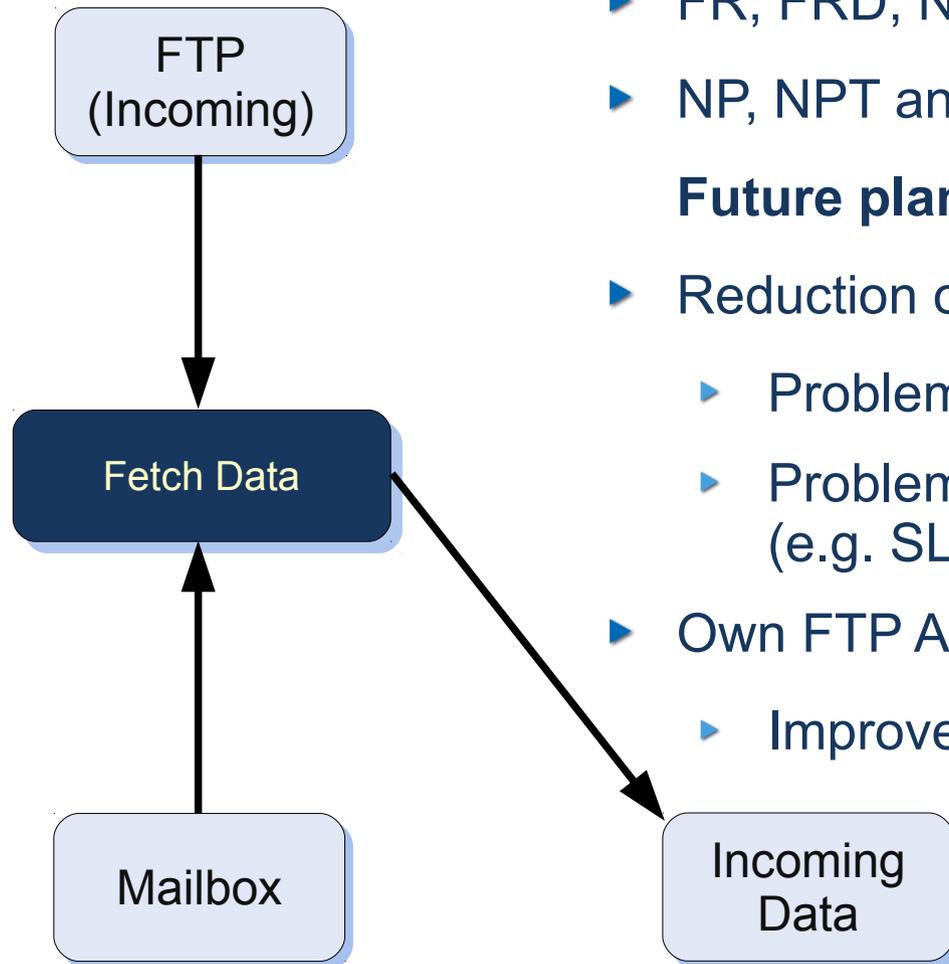
Past, Present:

▶ FR, FRD, NP, NPT and CPF via FTP

▶ NP, NPT and CPF via Mail

Future plan:

▶ Reduction of data transfer via mail

   ▸ Problems with junk mail

   ▸ Problems with encoding and attachments of mails (e.g. SLR-Mail)

▶ Own FTP Account for each station and provider

   ▸ Improvement of the security

C.Schwatke, Automated Data Management of SLR Data and Products at the EDC

Data Flow at EDC

```
FTP
(Incoming)
    |
    v
Fetch Data  ----> Incoming Data ----> Identify, Split and Check Data ----> Archive
    ^                                         |
    |                                         v
Mailbox                                   Database
```

FTP (Incoming)

Archive

Fetch Data

Identify, Split and Check Data

Mailbox

Incoming Data

Database
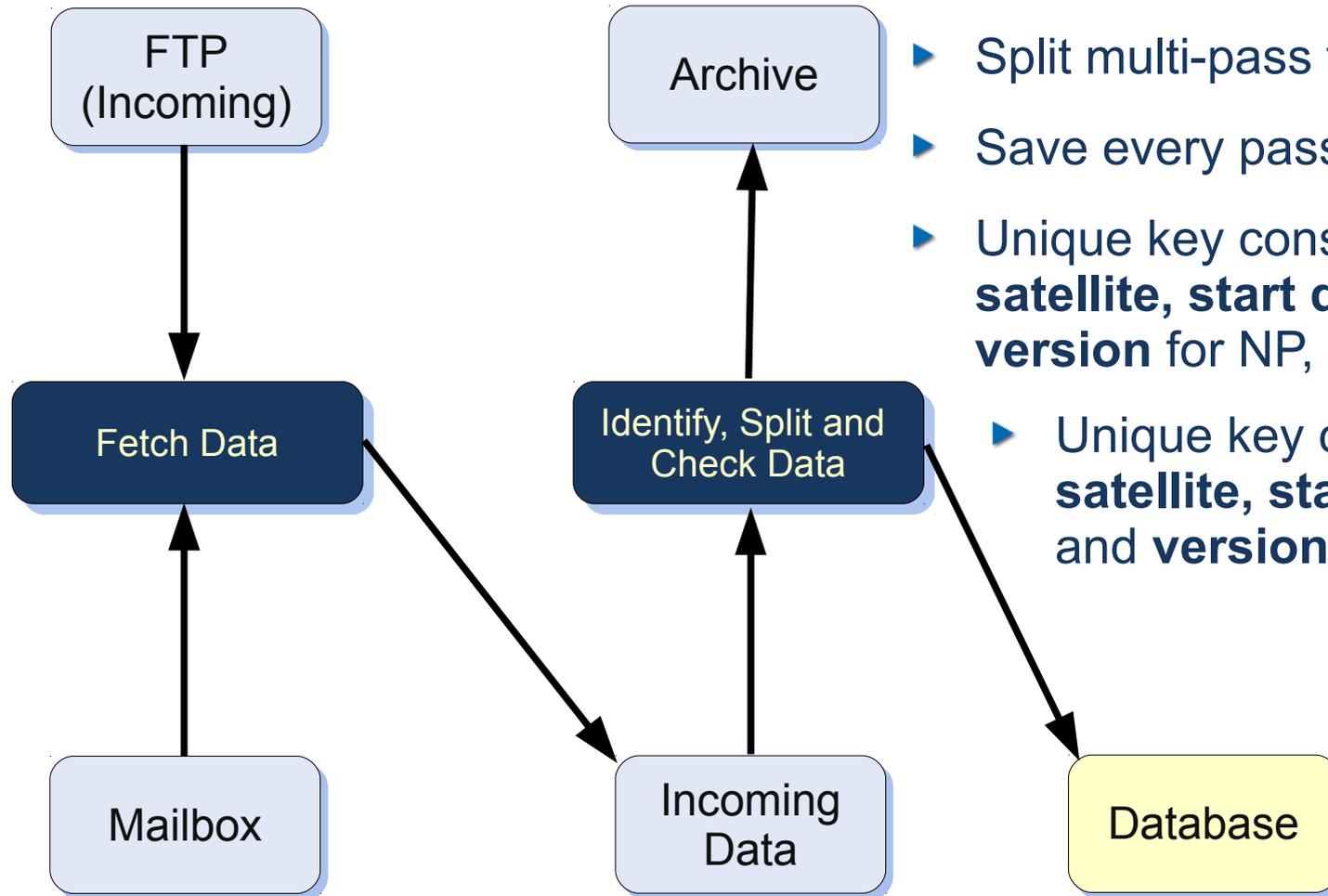
- ▶ Identification of data type (NP, NPT,...)

- ▶ Save original file/mail in archive

- ▶ Split multi-pass files

- ▶ Save every pass as new data set in DB

- ▶ Unique key consists of **station, satellite, start date, end date** and **version** for NP, NPT, FR, FRD data

  - ▶ Unique key consists of **provider, satellite, start date, end date** and **version** for CPF data

C.Schwatke, Automated Data Management of SLR Data and Products at the EDC

International Laser Ranging Service

DGFI

Data Flow at EDC

```
FTP
(Incoming)
    │
    ▼
Fetch Data ──────────┐
    ▲                │
    │                ▼
Mailbox          Incoming
                   Data
                     │
Archive              ▼
    ▲          Identify, Split and
    │          Check Data ──────────┐
    └──────────────┘                ▼
                                Database ⇄ Automatic
                                           Error Correction
```
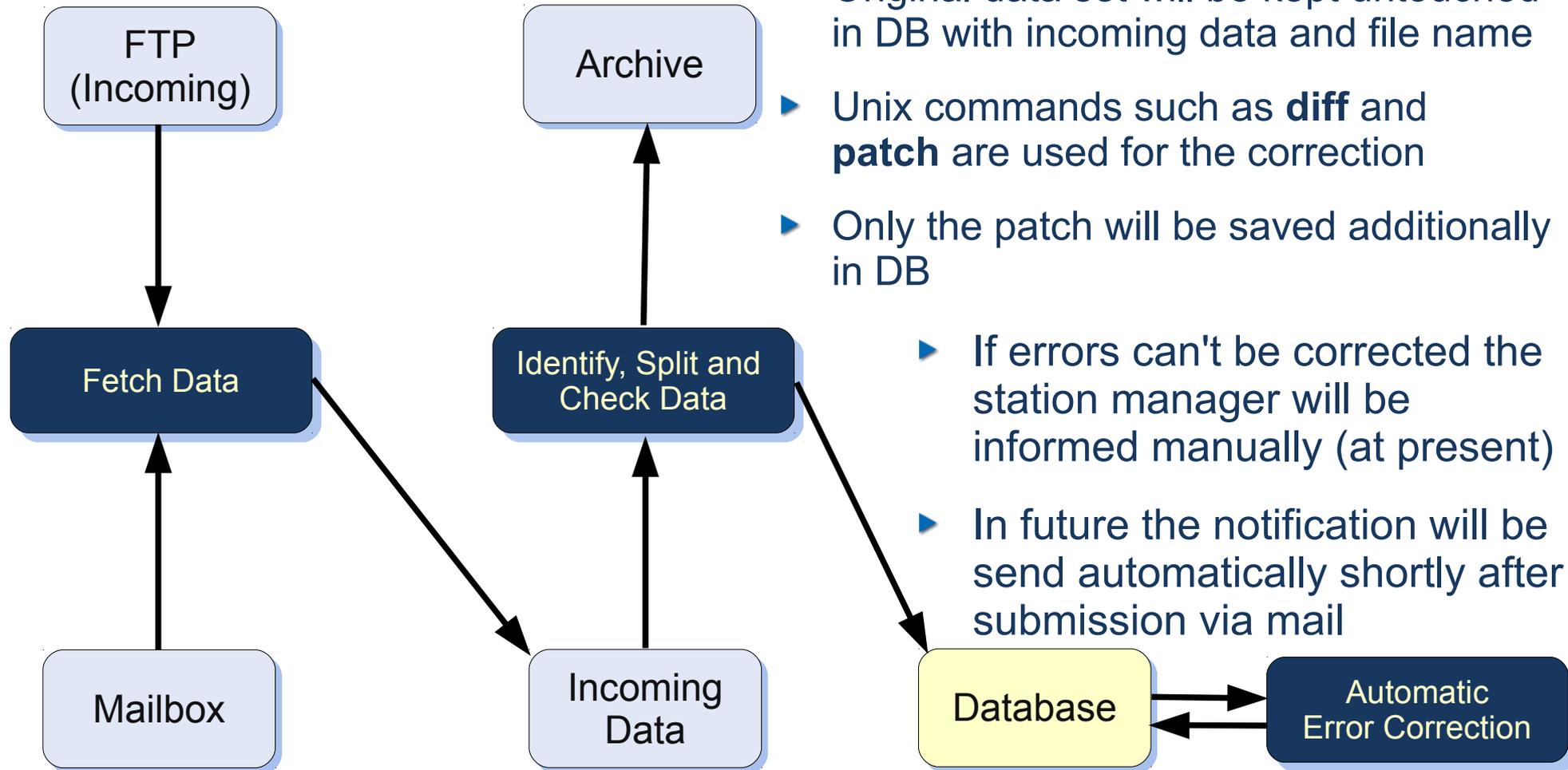
- Smaller format errors will be corrected automatically

- Original data set will be kept untouched in DB with incoming data and file name

- Unix commands such as **diff** and **patch** are used for the correction

- Only the patch will be saved additionally in DB

  - If errors can't be corrected the station manager will be informed manually (at present)

  - In future the notification will be send automatically shortly after submission via mail

Data Flow at EDC
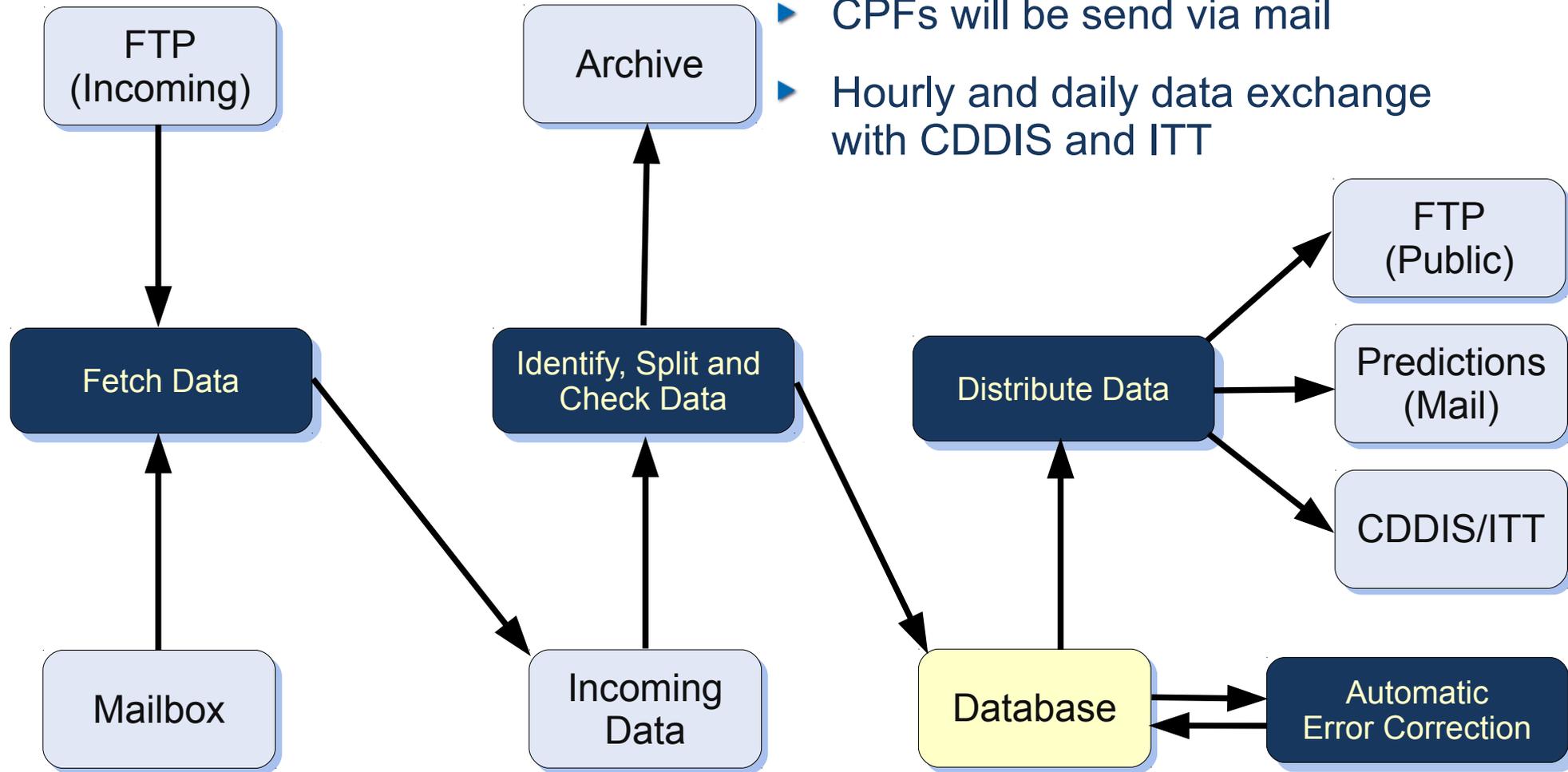
- All valid data sets will copied on FTP
- CPFs will be send via mail
- Hourly and daily data exchange with CDDIS and ITT

FTP (Incoming)

Archive

FTP (Public)

Predictions (Mail)

CDDIS/ITT

Fetch Data

Identify, Split and Check Data

Distribute Data

Mailbox

Incoming Data

Database

Automatic Error Correction

ILRS Mailing Lists

- The EDC maintains the **SLRMail**, **SLReport** and **Urgent** mailing list

- Currently every mailing list is handled by own scripts

- Problems:

  - Semi-automatic

  - Encoding of mails and attachments

  - Special tags for author and subject are needed

- Solution:

  - Transition to open source software Mailman

  - Automatic, no problems with encoding and special tags

  - Mailman is currently in testing phase at EDC

  - SLRMail will follow when transition is completed

C.Schwatke, Automated Data Management of SLR Data and Products at the EDC

EDC - Website

- New website of the EDC is under construction

- Features:

  - Near real time access

  - Current status of data sets (NP, NPT, FR, FRD and CPF) are available

  - Statistics about data holding

  - Direct access to operational database

  - Search function

- The launch will be announced via SLR-Mail in the next months

C.Schwatke, Automated Data Management of SLR Data and Products at the EDC