

**globalsafetyneedsglobalreliability**



**lima**  
lawful interception®

## Data Retention Challenges

 **group 2000**  
we develop telecom solutions

# Topics

- Introduction
- Challenges
- LIMA Data Retention



**globalsafetyneedsglobalreliability**

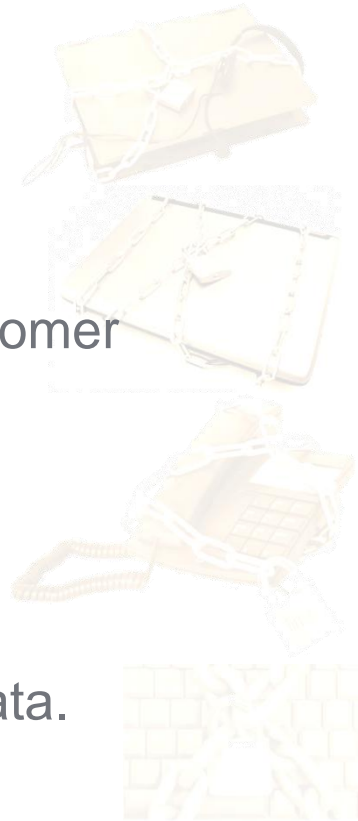
**lima**   
lawful interception®

# Introduction

 **group 2000**  
we develop telecom solutions

# The need for Data Retention

- European legislation requires CSPs to retain data.
- Retained data spans multiple domains: traffic data, customer data, location data.
- Retained data can span multiple type of networks.
- Access to the retained data is restricted.
- LEA requests must be serviced quickly, even for 'old' data.



## EU Directive

Data for identifying the following aspects has to be retained for a period of 6 to 24 months:

- the source of a communication,
- the destination of a communication,
- the date, time and duration of a communication,
- the type of communication,
- the user's communication equipment or what purports to be their equipment, and
- the location of mobile communication equipment.

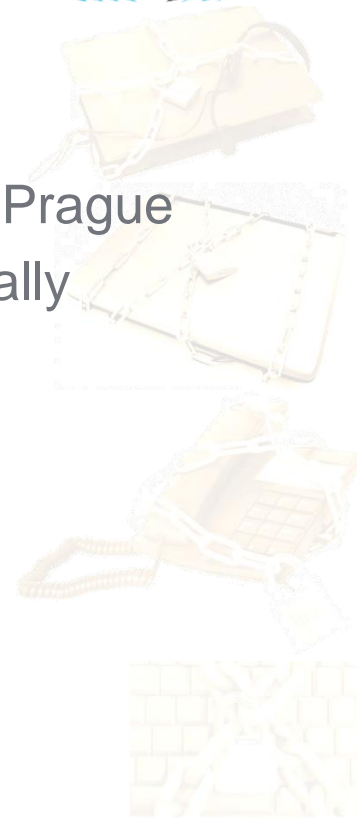
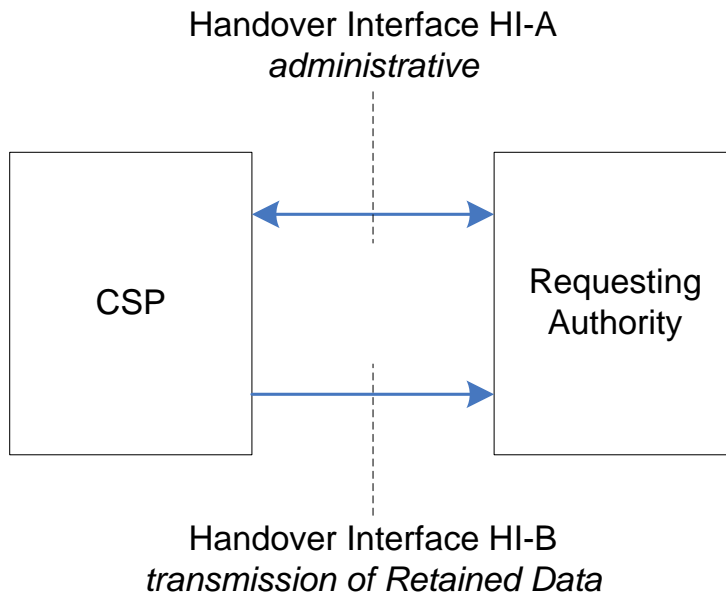




# Handover Interface

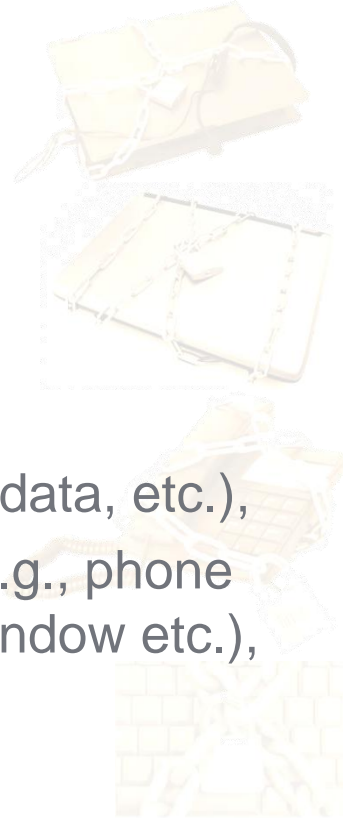
- ETSI is defining a Handover Interface

- Work-item is put forward for approval at TC LI 19 in Prague
- Both requests and replies are transferred electronically



# Handover Interface – request

- A requests contains of:
  - digital Signature (optional, for validation purposes),
  - CSP-ID as assigned to the Operator,
  - request-ID,
  - retained data category (e.g. subscriber data, usage data, etc.),
  - a set of identifications of the retained data subject (e.g., phone number, name, address, IMSI, time stamp or time window etc.),
  - requested period.



## DR life-cycle

The processes involved in Data Retention can be summarized as follows:

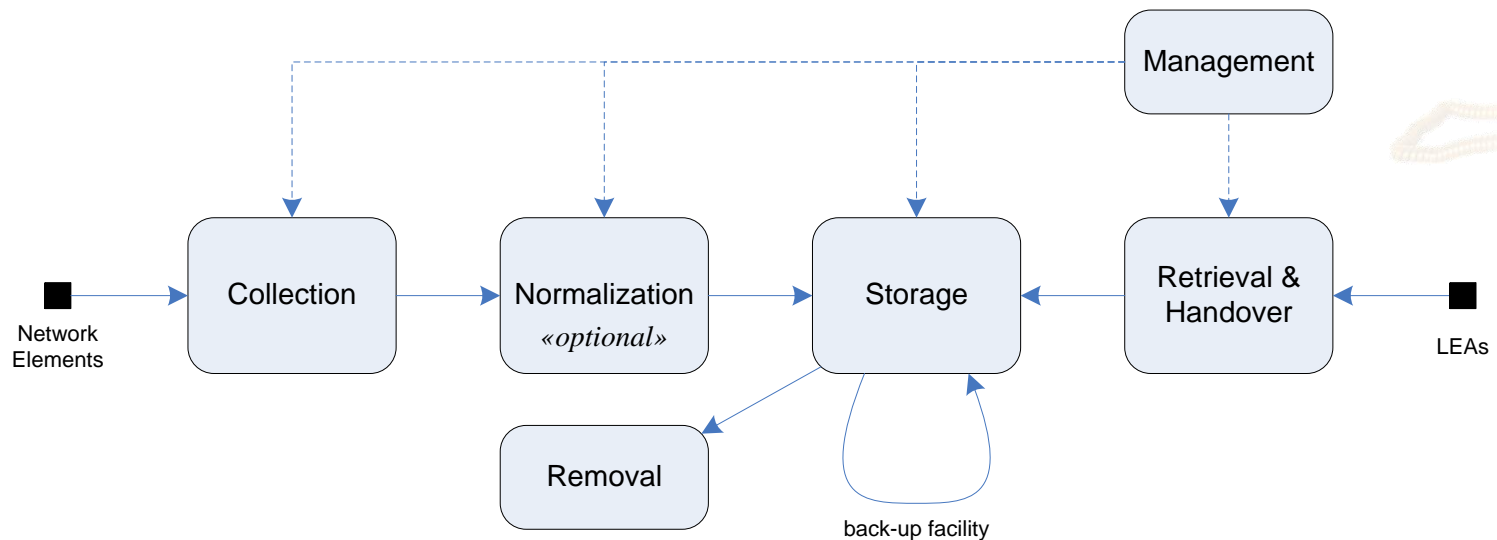
- *collect* the data to be retained,
- *prepare* the data for storage,
- *store* the data in a searchable way,
- *manage* LEAs and execute requests for Retained Data,
- *retrieve* the information from the Retained Data Store,
- *hand-over* the requested Retained Data to the LEA's premises, and
- *destroy* the data when the retention period elapses.





## DR life-cycle overview

- Preparation gives additional assurance that the right *amount* of data is retained and that the data meets the expected *quality*.
- Preparation is optional though...



**globalsafetyneedsglobalreliability**

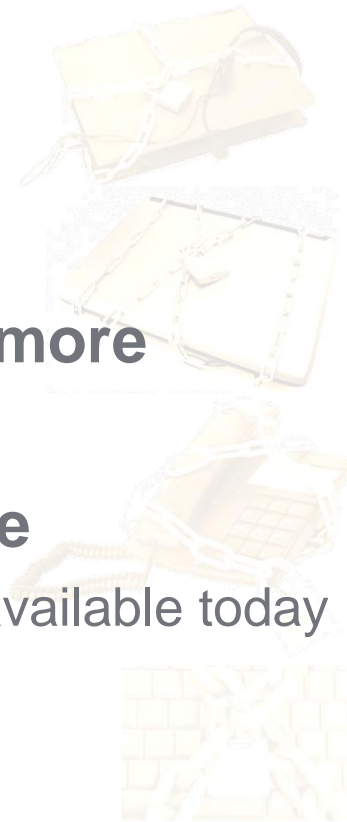
**lima**   
lawful interception<sup>®</sup>

# Challenges

 **group 2000**  
we develop telecom solutions

# Challenges

- **Millions of CDRs per day**  
*(dependent on CSP size and network type)*
- **Retention period of 6 months to 2 years or more**
- **Storage range approximately 10 – 100 TByte**
  - Compact storage systems of 48TByte in 4U height available today



## Organizational Impact

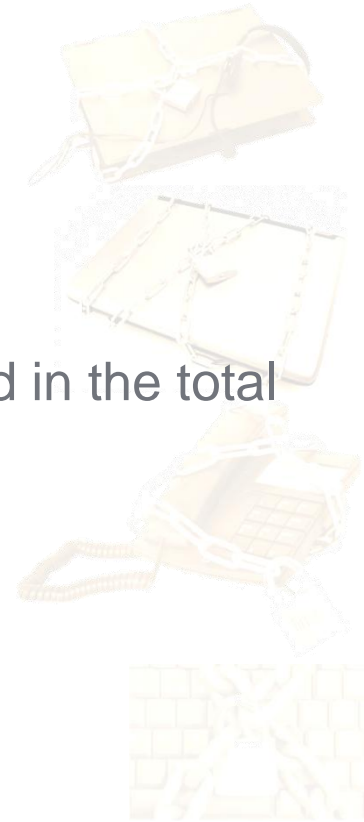
- Many requests for RD are expected
- Security budgets are typically not increased
- A high level of automation for handling RD requests is needed
- Single user interface for LI *and* DR is an advantage

More automation means less manpower...



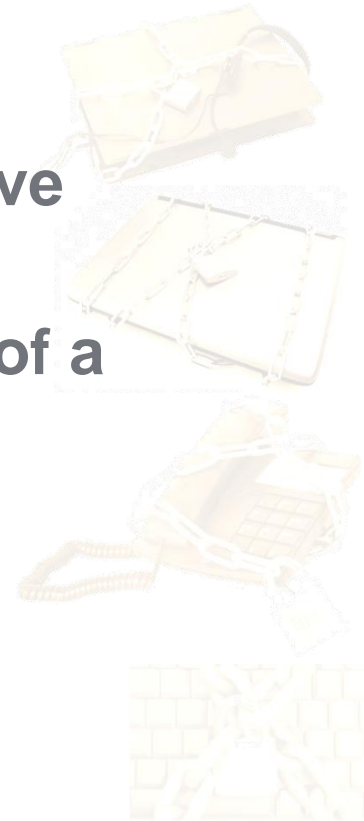
# Collection Phase

- **Deployment is non-standard**
  - Each CSP has its own mix of Network Elements
    - *Transport of information to RD site*
  - Various sources of information need to be integrated in the total DR infrastructure
    - *Dedicated conversions might be needed*



## Normalization Phase

- Data retrieved from various NEs tend to have different formats.
- Typically, only a subset of the information of a CDR has to be retained.
- Normalizing the data leads to:
  - Uniform data
  - Less data



# Retention Phase

## ■ Approach to Storage

- Relational database – *'heavy' weight*
- Object database – *does not perform well*
- Indexed files – *'light' weight*

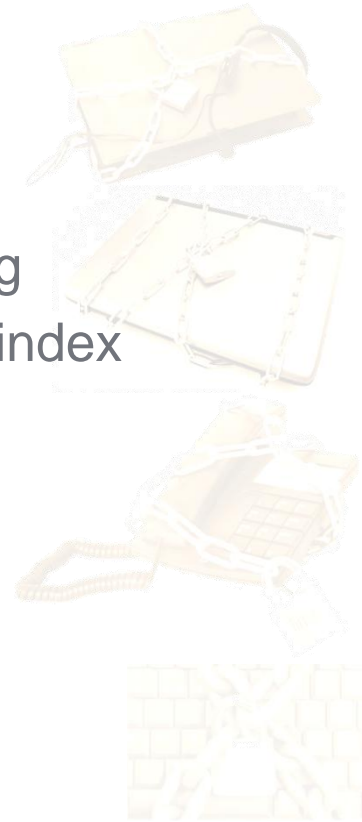
## ■ Huge amounts of information

- Scaling to hundreds of TByte
- Redundancy
- Back-up facilities



# Retrieval & Handover

- **Finding a needle in a haystack**
  - Efficient indexing many TBytes of data is challenging
  - Many indexing technologies need to re-arrange the index regularly
- **Handover Interface**
  - ETSI compliant
  - Support for country-specific extensions
  - Alternative interface (email, fax, CD/DVD)





## Conclusion – Functionality

- **Implementing the complete life-cycle leads to:**
  - High quality of the retained data
  - Highly automated processing of requests
- **Not implementing all phases leads to:**
  - Lower quality of the retained data (calculated risk)
  - Manual handling of many (or all) requests



## Conclusion – Investments

- Integration with CSP's network equipment
- Storage capacity is relatively cheap  
(less than €1000 / TByte)
- Software licenses are largest price-component
  - Licensing can be based on:
    - CDRs / day  
*software usage*
    - Number of subscribers  
*raises questions about transit-traffic and roaming subscribers*
    - Stored CDRs  
*hardware usage*
- Maintenance & Support fees



**globalsafetyneedsglobalreliability**

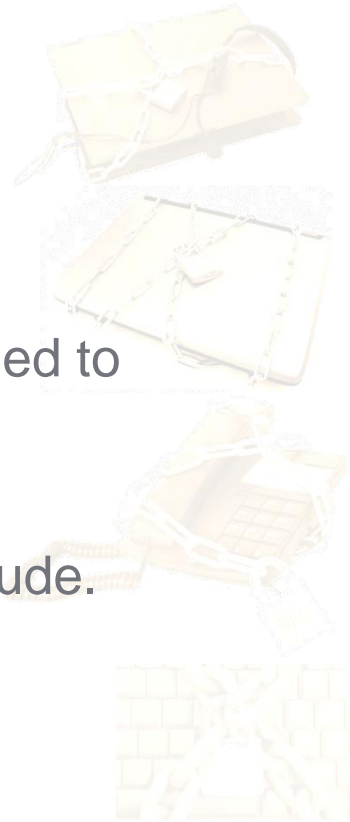
**lima**   
lawful interception®

# LIMA Data Retention

 **group 2000**  
we develop telecom solutions

# Facing the Obligation

- **Doing nothing is dangerous:**
  - When a request for retained data is received, you need to have information from the past!
  - “Wait until requested and then act” is a high-risk attitude.



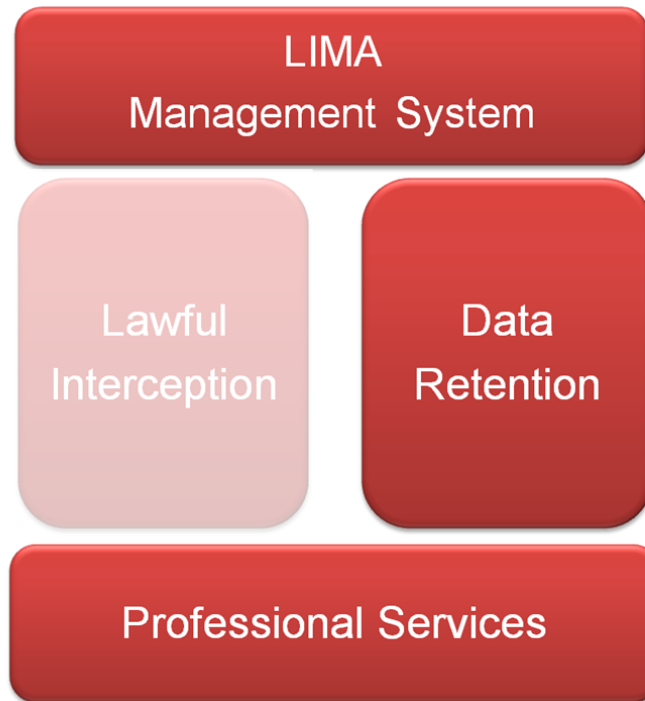
# Facing the Obligation

- **Characteristics of a Data Retention solution:**
  - Massively Scalable
  - Cost Effective
  - High Speed
  - Data Integrity Guaranteed
  - Workflow Management
  - Proven Solution

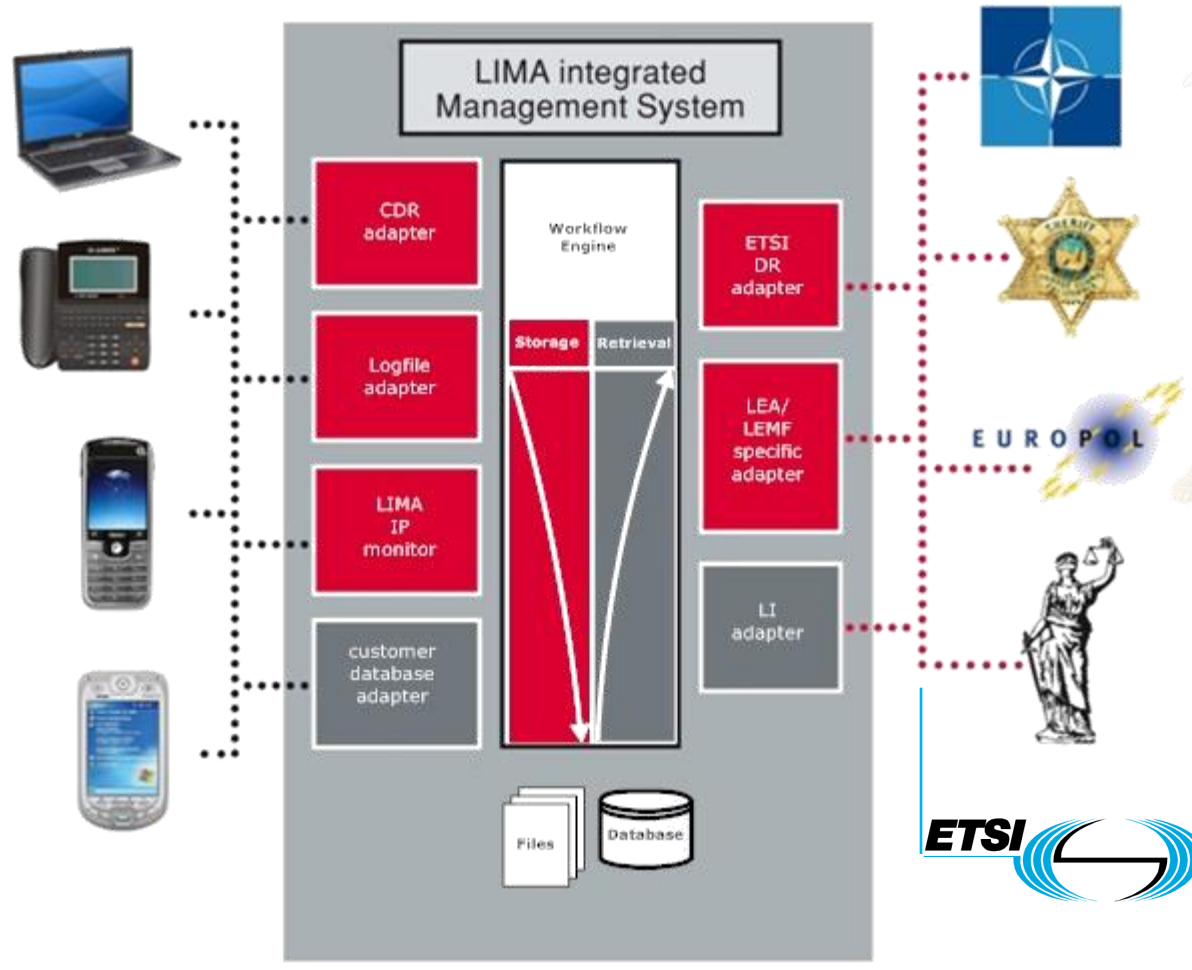


*DIY solutions can appear to be a cheap approach!*

# LIMA Integrated solution

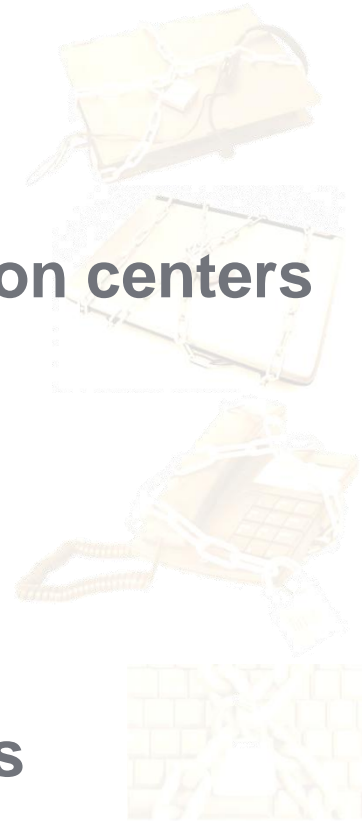


# LIMA Data Retention architecture



## Key Features

- **Internal process workflow configurable**
- **Web interface for CSPs and LEA interception centers**
  - Requests status control and alarms management
  - Centralized management of remote archives
  - New LEA services/requests configurability
- **Secure access to Retained Data**  
*(authentication, encryption)*
- **Integration API to support external systems**
- **Multi-operator architecture support**  
*(services bureau for CSPs)*





# Workflow Management

- **Customization of Data Retention process**
  - Warrant handling
  - Authorization steps
  - Data search
  - Presentation of data set
  - Cover page (*fax*)
  - Configuration of dispatch mechanism



# Input Adapters

## Collection and Normalization of data:

- Log file adapters
- Database adapters
- IP or SS7 probes
- LI Mediation devices
- Customer specific adapters



# Retention

- LIMA DR uses file based storage
  - High-performance indexing
  - Encryption
  - Compression



# Output Adapters

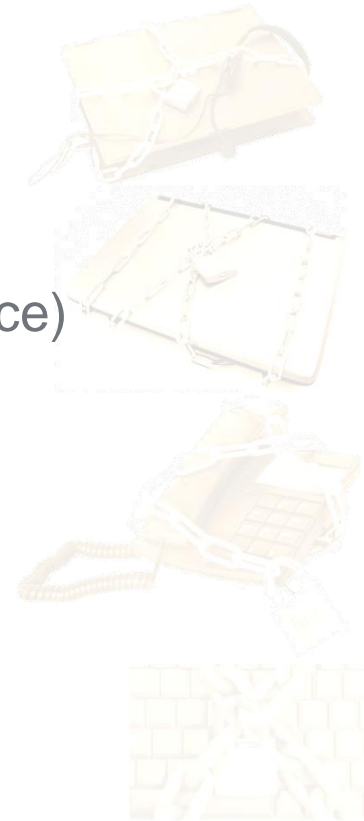
## Retrieval

- Efficiently search through retained data
- Presentation of data set
- Output via Fax, manual, ETSI DR



# System Integration

- **Integration into CSP network**
  - Each CSP has its own mix of Network Elements
  - Various sources of information (some already in place)
- **Configuration of data processing**
  - Information needs to be processed before storage
- **Operational aspects**
  - Workflow configuration
  - Training



# Lima Data Retention

- **Group2000's strength is in interfacing**
  - Partner handles storage & indexing the retained data
- **Integration of LI and DR in a single user interface**
- **Single vendor towards customer**
- **Hosting or Shared deployment for small CSPs possible**



**What  
can we  
do for  
you?**