

Data Classification and ILM

for a consistent

Storage Strategy

Josef Villa

Director Storage Solutions

S&T Group

- A prerequisite for a strategy are defined **metrics**.

- What are the taxonomies/attributes for a storage strategy ?
 - Confidentiality of data
 - Consistency of data
 - Continuity of business processes
 - Compliance with internal/external regulations

- As corresponding objectives (SLA´s) are defined in a very abstract and undifferentiated manner, they won´t become operational at all.

- Data offer attributes about technical structure:
 - Structured
 - Semi-structured
 - Unstructured
- Data own „aging“and „sourcing“ information
- But ! to develop a consistent storage strategy you need **additional** attributes:
 - Legal attributes such as retention period and audit proofing technology
 - Attributes for Continuity and Consistency
 - Attributes for Confidentiality of data

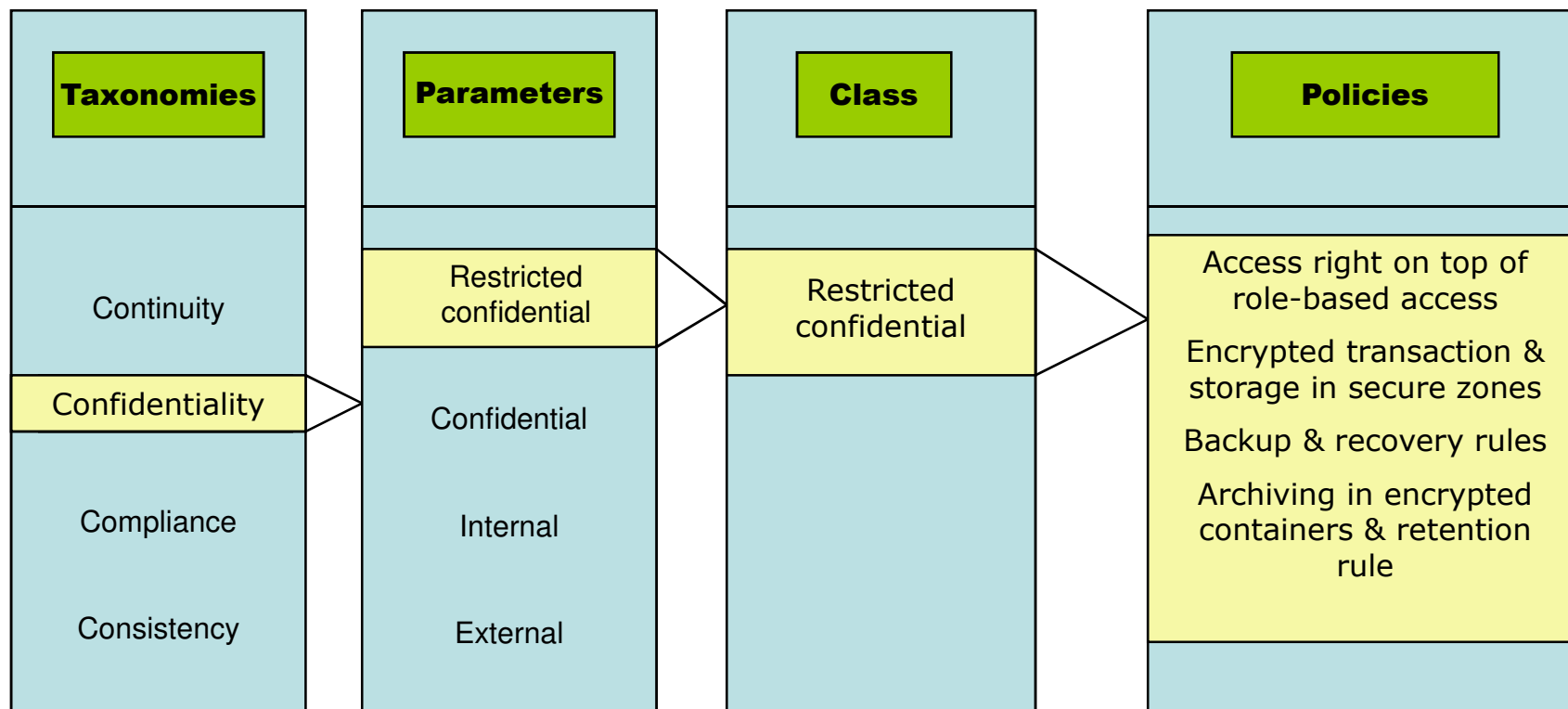
An Example of a Framework



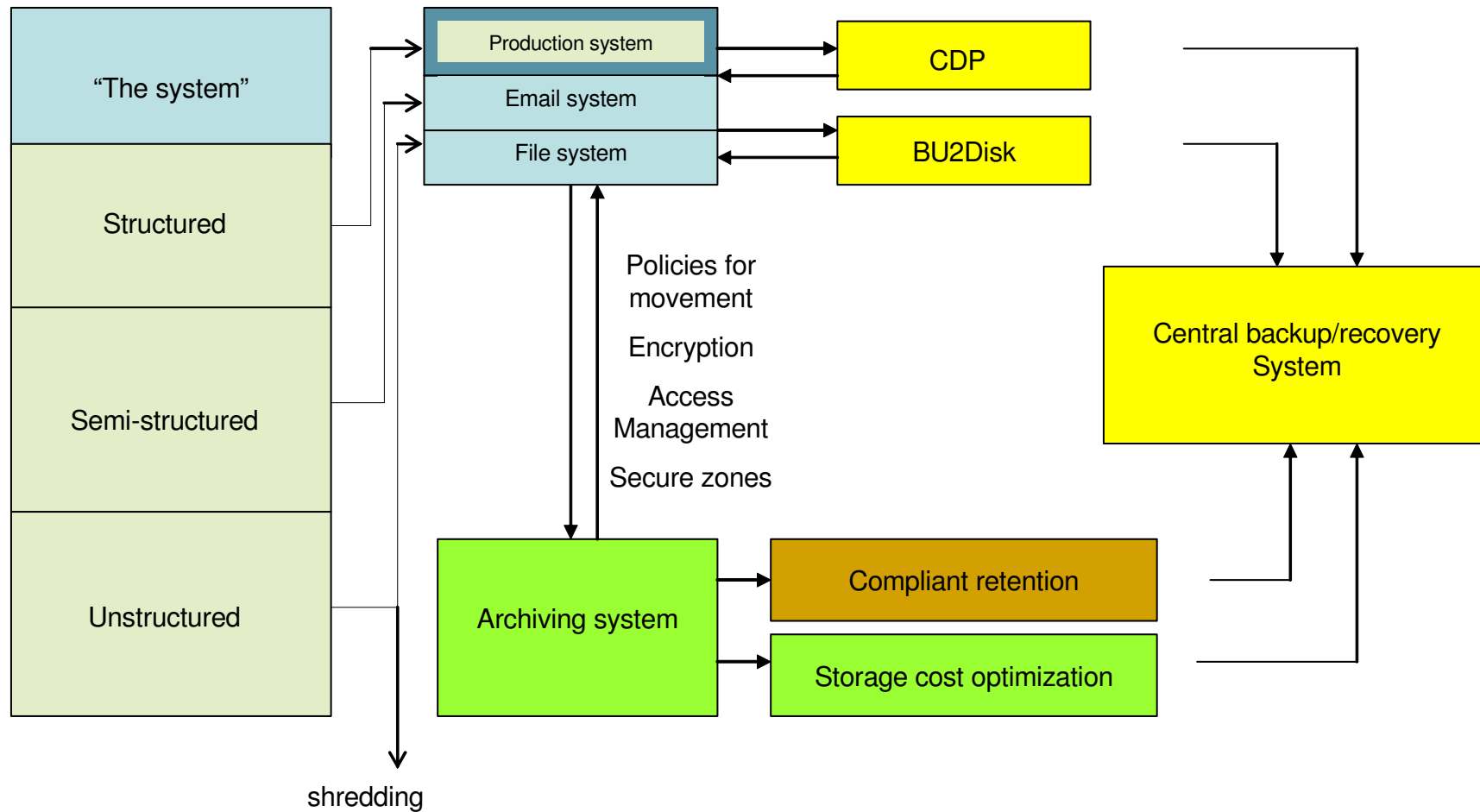
Attribute		Class	
Confidentiality	Strictly confidential	Confidential	Internal
Continuity	Critical = 99,99%	Sensitive = 98%	Non-Critical = 85%
Consistency/RPO	Critical = 3 Minutes	Sensitive = 2 Hours	Non-Critical = 3 Days
Consistency/RTO	1 Hour	1 Day	3 Days
Retention Period	3 Years	7 Years	15 Years
Audit Proof Tec	yes	no	

- Choose a **methodical** template based **approach** together with the data owner
- **Complexity** of the model is a function the number of axes and granularity of parameters (5-10)
- **Legal requirements** and **compliance issues** should be known and included
- Make the **relation of metrics** and cost very clear!

Data Classification Method



Scenario: a Strategy of functional Differentiation



Protective Measures for each Class

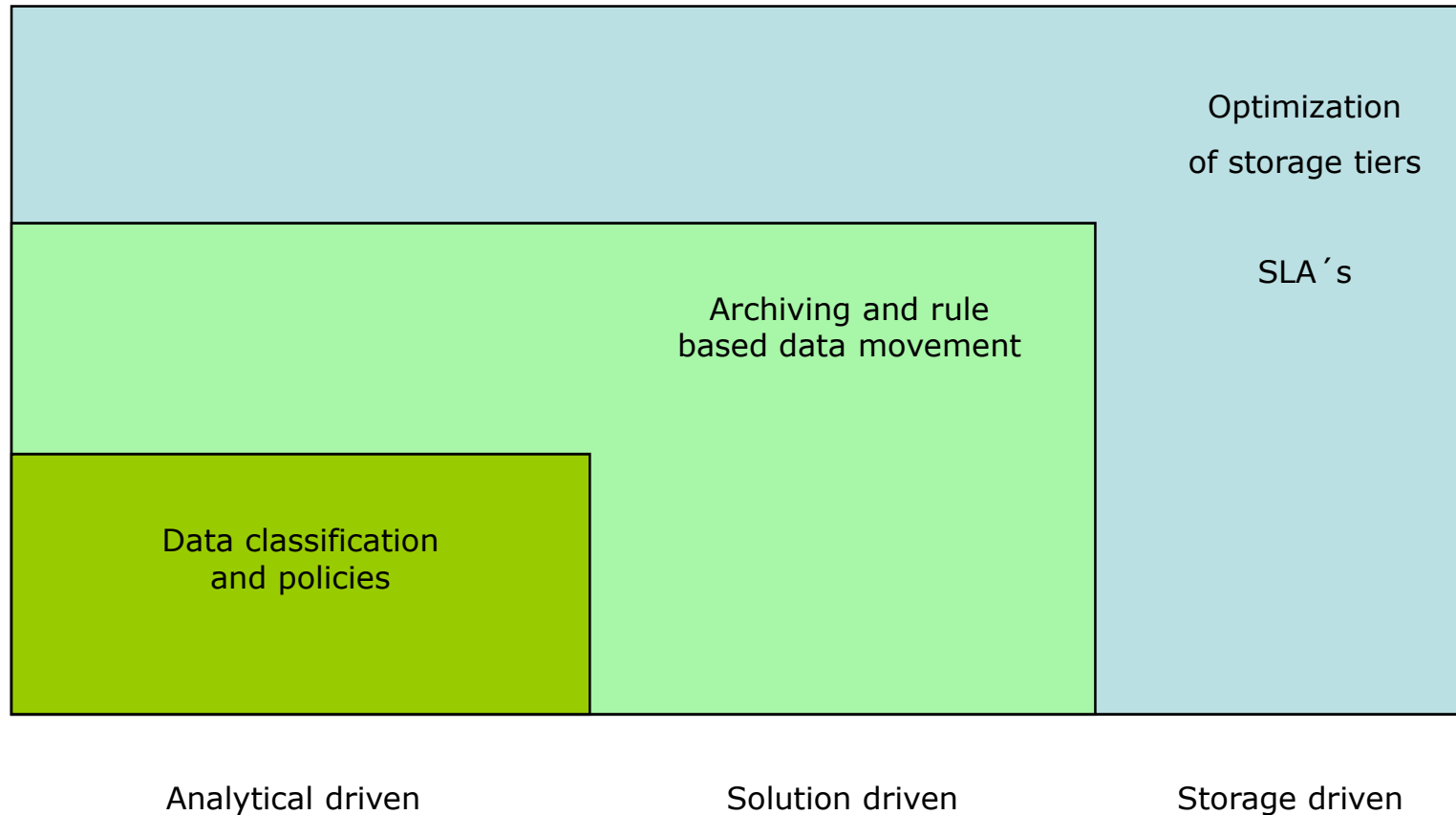


Class	Protective measures
Restricted confidential	<p>Authentication: 2-factor/strong password, excluded from role based rights-management</p> <p>Application: timeout/short term screen saver/PC security/ Security updates/disabled utilities/SIEM log management & access audits/disabled devices/IP wrap</p> <p>Network: certified devices/encryption for remote/firewall/IDS</p> <p>Storage: secure zoning/encrypted container for archiving</p> <p>Forwarding: pre-defined addressees, non-disclosure</p> <p>Recovery under unchanged access rights</p> <p>Signed document of users</p>
Confidential	<p>Authentication: single factor/strong password</p> <p>Application: security updates, disabled utilities/devices</p> <p>Network: firewall, IDS/SIEM & audits</p> <p>Forwarding: non-disclosure training</p>
Internal	<p>Passwords/security updates/audits</p>

Policy relevance

Data confidentiality	Embedded in a complete Security Policy Data shredding policy
Data recovery	Backup & Restore policies under confidentiality constraints
Data availability	Risk & impact analysis of business critical processes Business Continuity Plan Disaster Recovery Plan
Data retention	Policies for archiving under compliance attributes. Data deletion policy !!!

The big picture on ILM

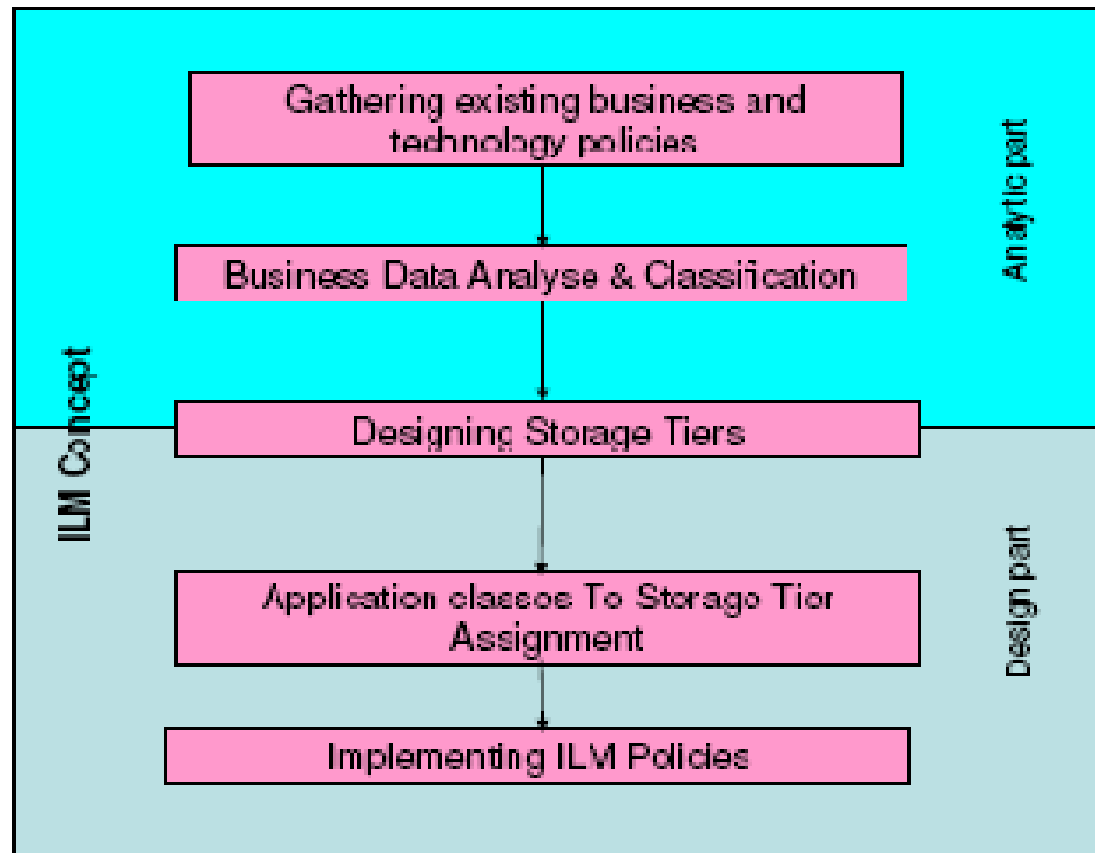


Information Life Cycle (ILM) is a corporate management system, based on changing business value of data and policies to move them.

a holistic approach to:

- migrate different data classes to appropriate structures
- define SLA's and policies, including legal & compliance
- define archiving and retrieval by policies
- build the system with solutions and optimized platforms
- splitting data into “three worlds” with intelligent exchange rules
- minimize backup/recovery times and improve performance

The steps



Service Level Catalog – Client Sample



	Alignment Attributes		Tier 1	Tier 2	Tier 3
	Scheme	Specification			
Primary Storage	Guaranteed Performance	Performance throughput per port (I/O sec)	5,000+	Up to 3,500	Up to 1,500
		Response time (ms)	< 8ms	12-30ms	12-30ms
	Availability	Maximum unplanned downtime per year (hours)	< 1	< 18	< 44
Archiving Storage	Confidentiality	Secure zoning	yes	no	no
		Encrypted containers	yes	yes	no
	Availability	Maximum downtime (hrs)	< 1	< 18	< 180
	Retention & Disposition	Retention period	< 30 years	< 10 years	< 3 years
		Data shredding compliance	Yes	No	No
	Discovery	Access	immediately	4 hours	48 hours
	Data Integrity	Guarantee of authenticity	Yes	No	No
Offsite	Recovery point objective	< 1 minute	< 24 hours	< 72 hours	
Operational Recovery	Recovery Classification	Recovery classification	Complete application restore	Complete application restore	File or file system restore
	Operational Recovery Point Objective	Amount of data loss	none	2 hours	24 hours
	Operational Recovery Time Objective	Time required for recovery	< 30 minutes	< 120 minutes	7 GB/minute
	Recoverability	Ability to recover backed up data	100%	100%	98%
	Retention period	Length of time that data is retained	minutes	4 hours	3 Weeks

- Keep this simple model simple
- Define the classes together with key users (“data owner”)
- Show relation of classes to a) regulations and to b) cost
- introduce classification tools for unstructured data to meet privacy protection and retention policies
- Inform employees about defined standards and behavior
- Data classification delivers the information to start optimizing your storage infrastructure, fulfilling given constraints

Splitting data into “worlds” instead of divisional “silos”:

productive, archive & trash*

- Applying different priorities, solutions and technologies
- Alignment on business objectives
- Policies for data movement according to business value
- Breaking the data growth for production systems
- Reducing back-up and recovery windows
- Managing compliance in an auditable way

* A data deletion/shredding policy would complete this chapter.

The outcome creates value.

Do you agree ?

