02 infra TI

Armazenamento de Informações

Universidade Presbiteriana Mackenzie

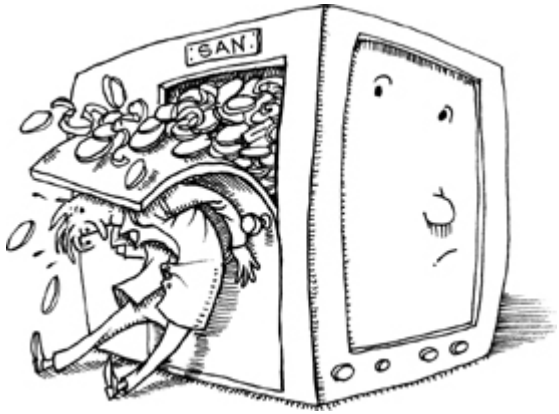
Crescimento dos dados e da importância das informações; Tipos de dados; Evolução das tecnologias de armazenamento; Estrutura e requisitos do data center; Ciclo de Vida da Informação;

Informação e dados

Informação : cada vez mais importante

Crescimento exponencial da importância, do volume e da dependência do mundo corporativo por informações

Aumentam, portanto, os desafios relacionados à proteção e ao gerenciamento dos dados

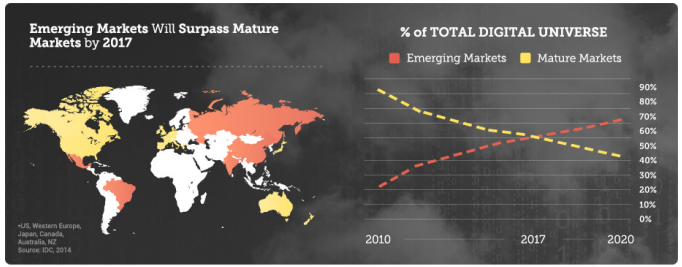


Crescimento exponencial

http://www.computerworld.com/s/article/9217988/World_s_data_will_grow_by_50X_in_next_decade_IDC_study_predicts

Computerworld - In 2011 alone, 1.8 zettabytes (or 1.8 trillion gigabytes) of data will be created, the equivalent to every U.S. citizen writing 3 tweets per minute for 26,976 years. And over the next decade, the number of servers managing the world's data stores will grow by ten times.

http://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm



Ooops a break: KB, MB, GB, :-) B

1 Megabyte =
1 million bytes
a tablespoon of sand

1 Gigabyte =
1 billion bytes
patch of sand—
9”square, 1’ deep

1 Terabyte =
1 trillion bytes
a sandbox—
24’ square, 1’ deep

1 Petabyte =
1,000 terabytes
a mile long beach—
100’ wide , 1’ deep

1 Exabyte =
1,000 petabytes
the same beach—
from Maine to North Carolina

1 Zetabyte =
1,000 exabytes
the same beach—
along the entire US coast

1 Yottabyte =
1,000 zetabytes
enough info to bury the entire US under 296 feet of sand

Crescimento: exemplo 1

10,000,000,000 photos

2-3 Terabytes of photos are being uploaded to the site every day

Serve over 15 billion photo images per day

Photo traffic now peaks at over 300,000 images served per second

Crescimento: exemplo 2

Inglaterra: Uma câmera de vigilância para cada 14 cidadãos
4 milhões de câmeras registrando imagens diariamente

Você tem ou pode encontrar outros exemplos na Internet ?

O desafio do armazenamento:
Armazenar, proteger, otimizar e influenciar essa enorme quantidade crescente de dados

Desafio

O desafio do armazenamento:

Armazenar, proteger, otimizar e influir* nessa enorme quantidade crescente de dados

influir*, pense em como o armazenamento suporta a capacidade de gerar informações sobre os dados

Tipos de dados

Estruturados
X
Não Estruturados

→

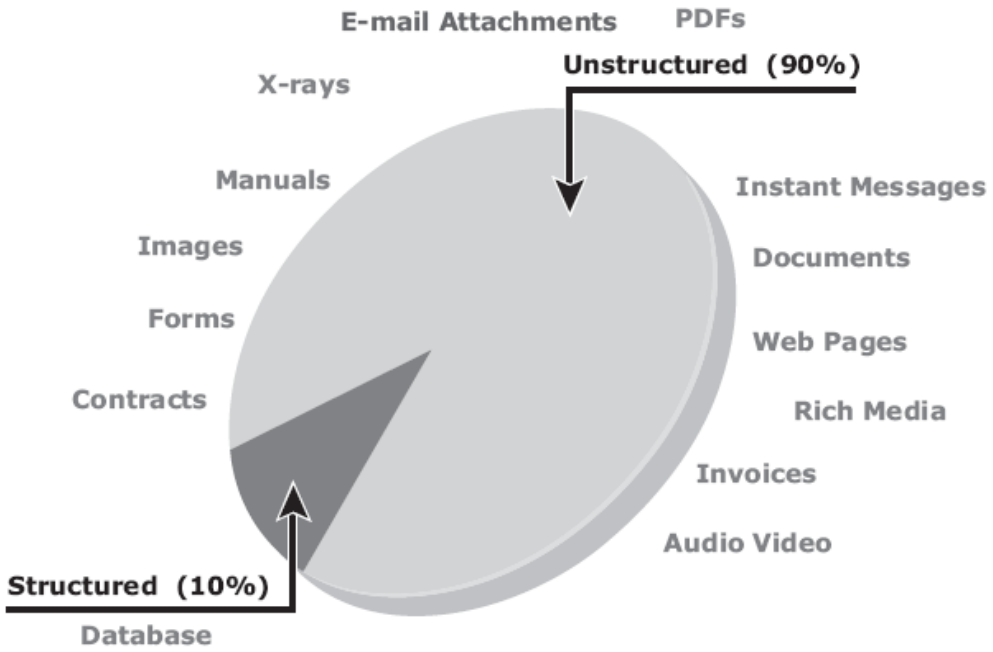


Figure 1-2: Types of data

Big Data: novos desafios

Big Data:

Novos desafios para armazenamento de dados nos centros de informação

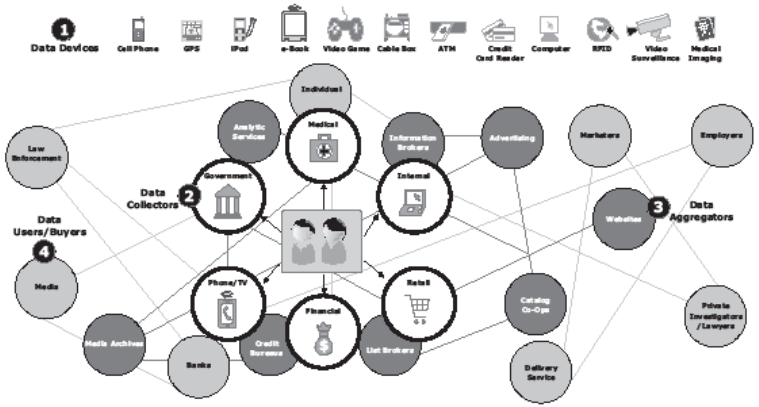


Figure 1-3: Big data ecosystem

Dispositivos de armazenamento

Os dispositivos de armazenamento variam conforme o tipo de dados, a velocidade com que esses são criados e usados, e a capacidade.

Devices, such as a media card in a cell phone or digital camera, DVDs, CD-ROMs, and disk drives in personal computers are examples of storage devices.

Businesses have several options available for storing data, including internal hard disks, external disk arrays, and tapes.

Evolução

Evolução dos dispositivos de Armazenamento:

Do armazenamento interno não inteligente para o armazenamento em rede inteligente.

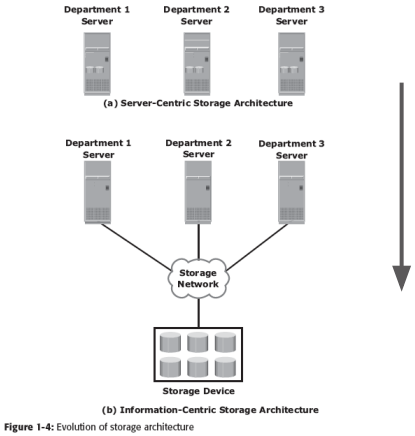
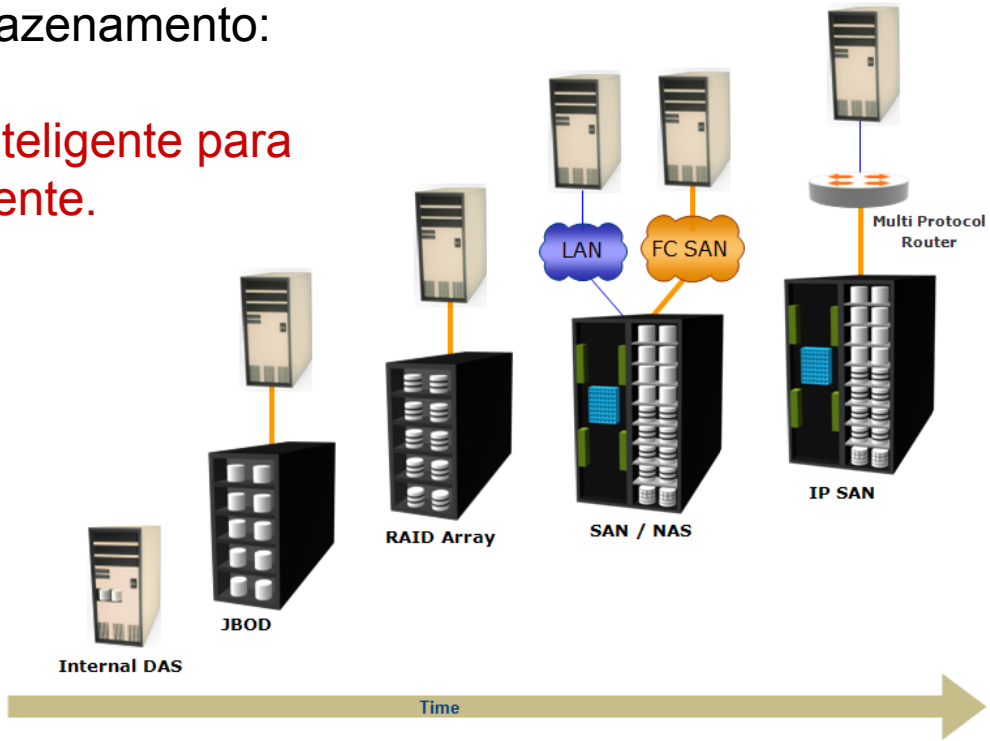


Figure 1-4: Evolution of storage architecture



Evolução

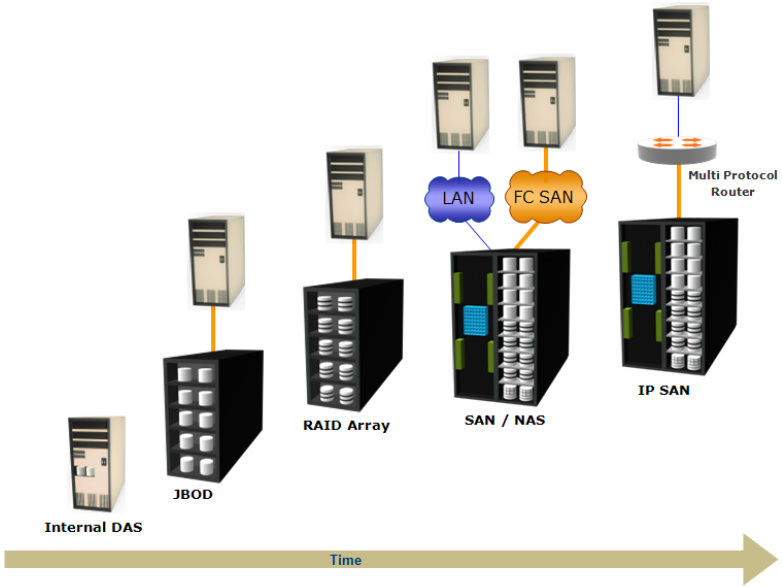
Redundant Array of Independent Disks (RAID)

Direct-attached storage (DAS)

Storage area network (SAN) This is a dedicated, high-performance Fibre Channel (FC) network to facilitate block-level communication between servers and storage.

Network-attached storage (NAS) This is dedicated storage for file serving applications. Unlike a SAN, it connects to an existing communication network (LAN) and provides file access to heterogeneous clients.

Internet Protocol SAN (IP-SAN) One of the latest evolutions in storage architecture, IP-SAN is a convergence of technologies used in SAN and NAS.



Data center: 5 Componentes Chave

Application: A computer program that provides the logic for computing operations

Database management system (DBMS): Provides a structured way to store data in logically organized tables that are interrelated

Host or compute: A computing platform (hardware, firmware, and software) that runs applications and databases

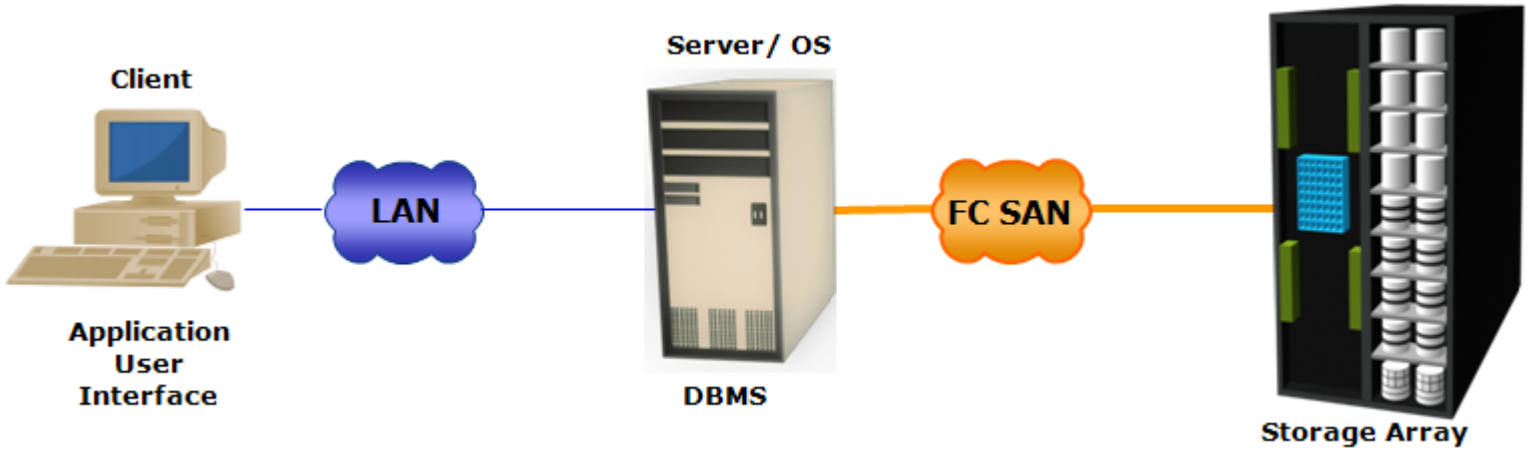
Network: A data path that facilitates communication among various networked devices

Storage: A device that stores data persistently for subsequent use.

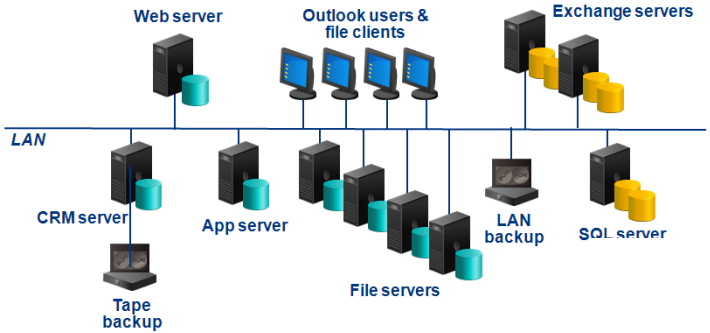
These core elements are typically viewed and managed as separate entities, but all the elements must work together to address data-processing requirements.

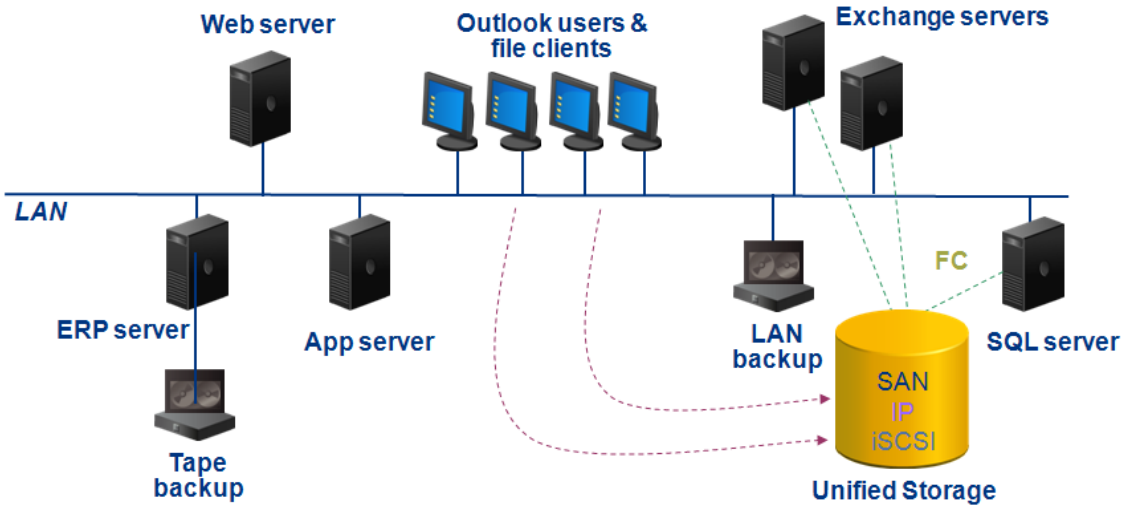
Arquitetura típica

Uma arquitetura típica de processamento de um data center usando uma rede de armazenamento (SAN) em um data center



Old to Modern approach





Características Chave de um DC

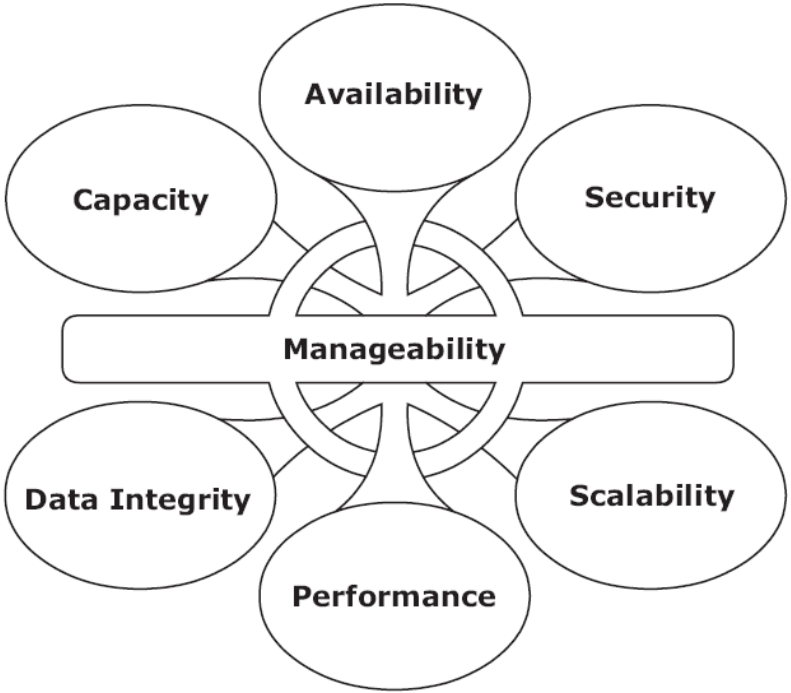
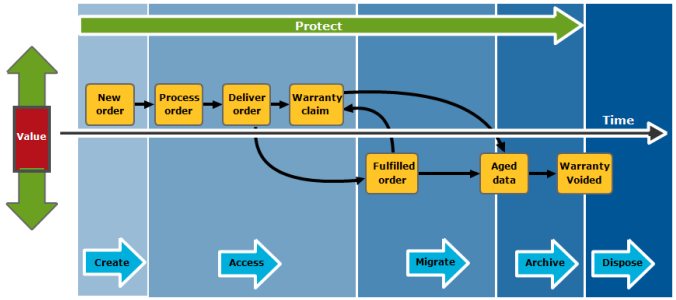


Figure 1-6: Key characteristics of data center elements

ILM Information Life Cycle Management

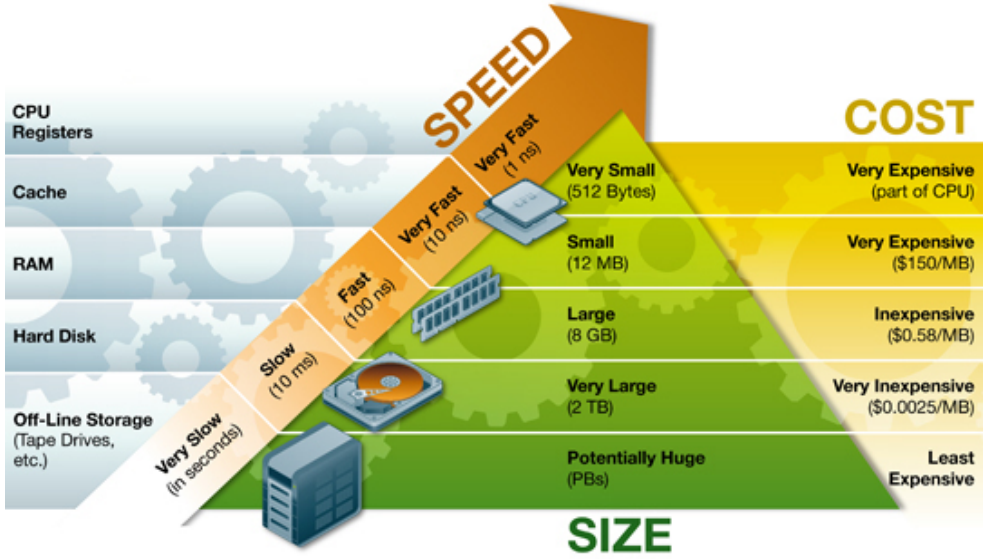
The information lifecycle is the “change in the value of information” over time.
When data is first created, it often has the highest value and is used frequently. As data ages, it is accessed less frequently and is of less value to the organization. Understanding the information lifecycle helps to deploy appropriate storage infrastructure, according to the changing value of information.

A proactive strategy that enables an IT organization to effectively manage the data throughout its lifecycle

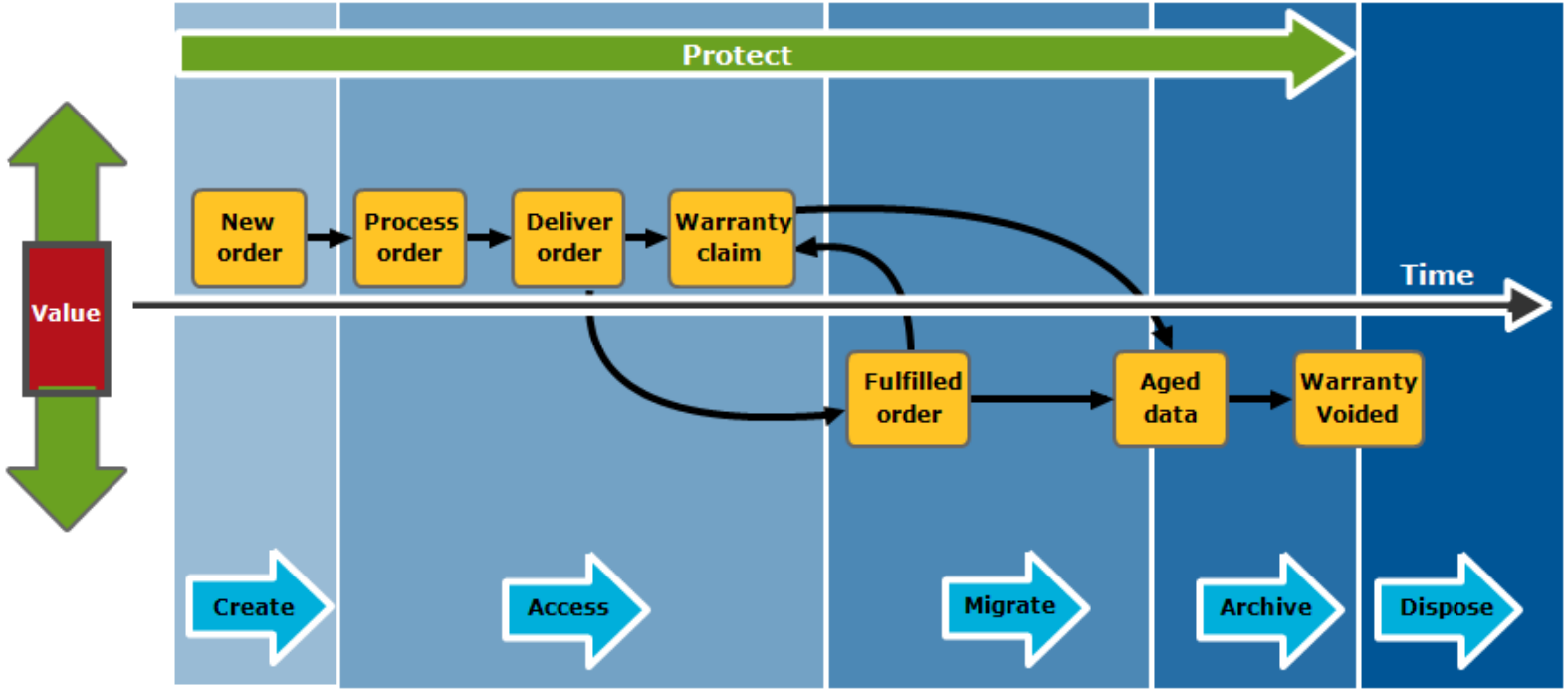


ILM Storage Hierarchy

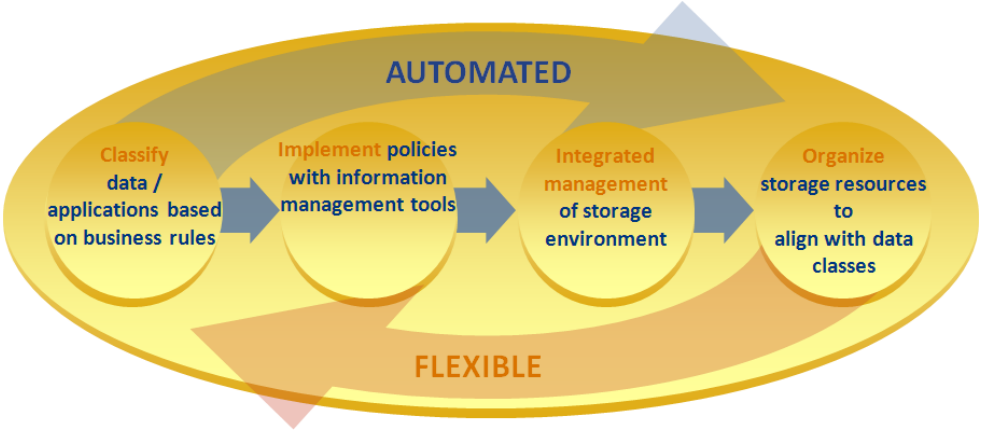
Uma ideia básica é que Custo x Velocidade x Capacidade de Armazenamento definem naturalmente Tiers de Armazenamento. A informação, para ser armazenada de forma eficiente, precisa ter um custo de armazenamento correspondente ao seu valor para a Organização



ILM Information Life Cycle Management



ILM Process



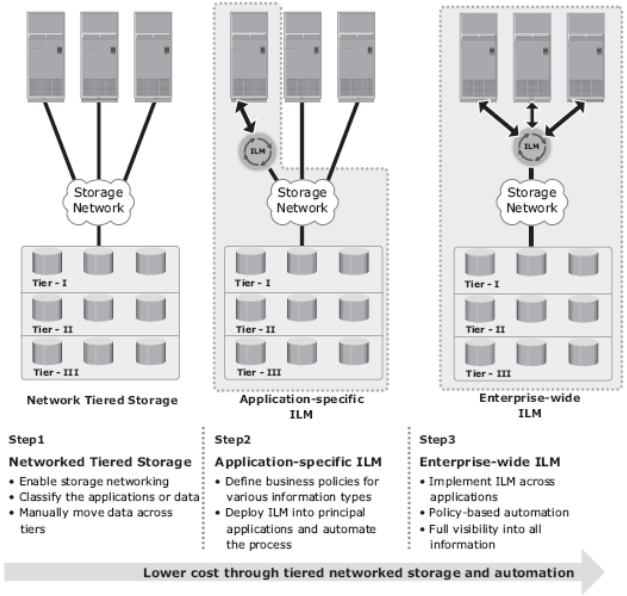


Figure 1-8: Implementation of ILM

ILM Benefícios

Improved utilization

Tiered storage platforms ← Low Costs

Simplified management

Processes, tools and automation

Mas há um custo $ e, na prática, nem sempre é algo fácil de implementar um ILM de modo eficiente

Simplified backup and recovery

A wider range of options to balance the need for business continuity

Maintaining compliance

Knowledge of what data needs to be protected for what length of time

Lower Total Cost of Ownership

By aligning the infrastructure and management costs with information value

Discussão e exercícios

Um crescimento exponencial dos dados e dos Data Centers pode significar um incremento igual de profissionais e recursos ($) em TI nos próximos anos?

Considere os dados de um venda no caixa de um supermercado. O valor dessa informação é o mesmo ao longo do tempo (primeiros dias, meses e após um ano por exemplo)?

Cite facilidades ou recursos que você espera de uma ferramenta de automação de ILM.

Na sua opinião que tipo de dado, estruturado ou não estruturado, parece ter um crescimento maior hoje e por que?

Que vantagens você vê no armazenamento em rede sobre o interno?

Leitura recomendada

Capítulo 1

Information Storage and Management Storing, Managing, and Protecting Digital Information in Classic, Virtualized, and Cloud Environments

2nd Edition Edited by Somasundaram Gnanasundaram, Alok Shrivastava