Overview

June 2005

intermine

FileCensus

Instant Storage Insight

Agenda

Objective - To demonstrate how FileCensus helps large enterprises manage their unstructured data across complex heterogeneous server and storage networks using intelligent ILM - Information Lifecycle Management.

Company Overview

- Incorporated in 1999
- Founder - ex-Novell
- ILM has always been major company focus

- US Headquarters in San Francisco
- Distributors in Japan & Europe
- Development Centre in Canberra, Australia.

- Major partnerships with Network Appliance and HP
- Excellent Customer Base

Intermine products have an excellent proven track record with Fortune 1000 companies

intermine FileCensus
Instant Storage Insight

Complexities of ILM

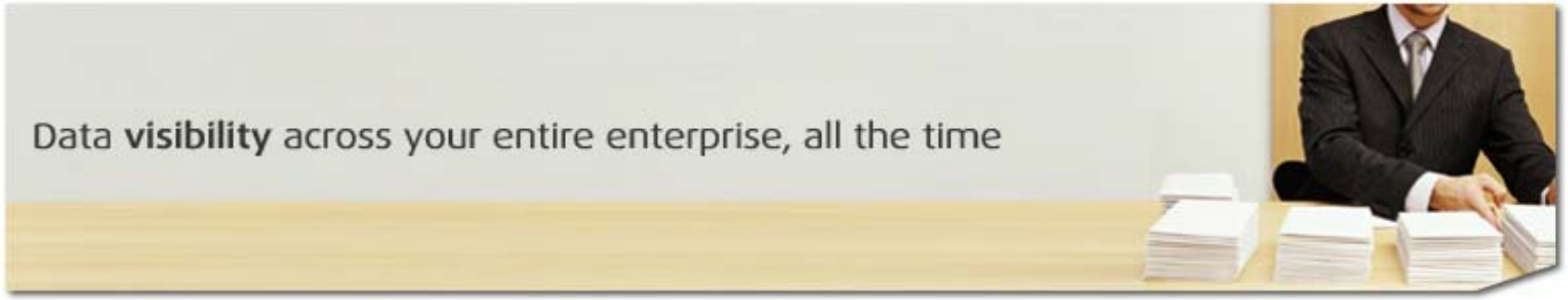
- Exploding information growth > 60% per year
- Limited budgets < 10% growth creating huge gap!
- Information - Strategic Value/ Dynamic Nature/ Always Ageing
  - No longer an infrastructure problem
  - Limited resources and valuation systems
- Compliance regulation is also driving SRM projects
  - CAGR 2003-2006 Forecast – 64%
  - US & UK 2004 deadlines

ILM Planning

- Understanding the information on your network is the **first step** to implementing a SRM/ILM strategy.

- Once you discover what is happening with the information on your network you can start examining and adopting best practices and implement your own specific policies.

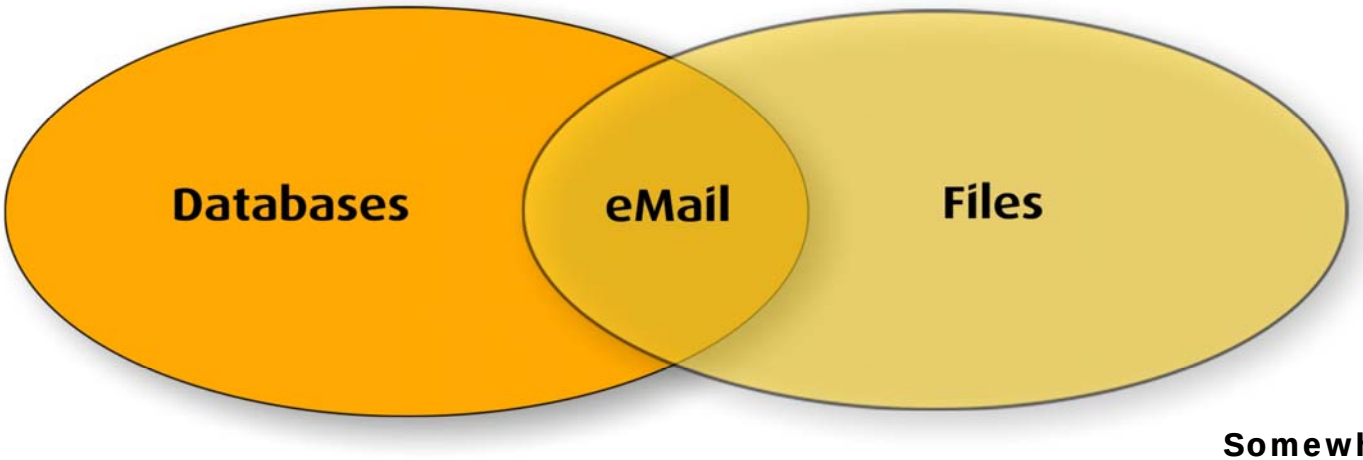
- Knowing and using this information are two different things
  - Top 5 Database Applications, Email, Other?



intermine FileCensus
Instant Storage Insight

FileCensus - Areas of Relevance

Structured vs. Unstructured



Compliance Problem Area

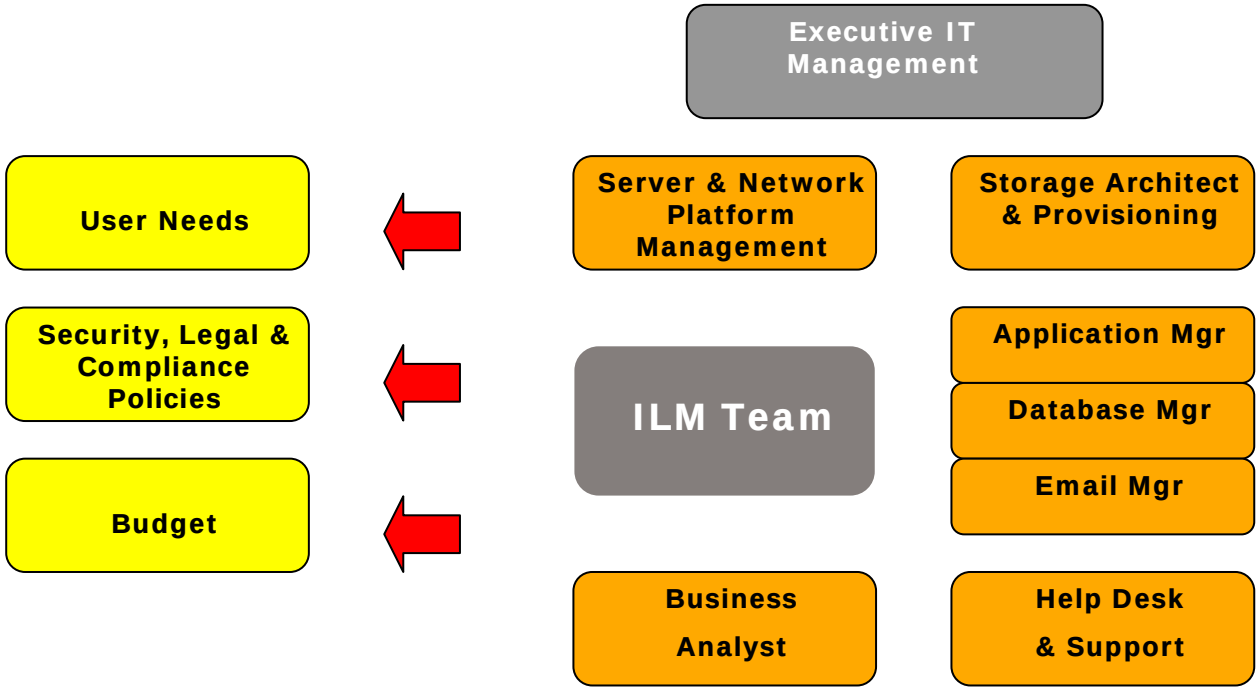
Major Growth Area

Somewhat Unmanaged

DBAs	Email Mgrs	Who?
35%	**25%**	**40%**

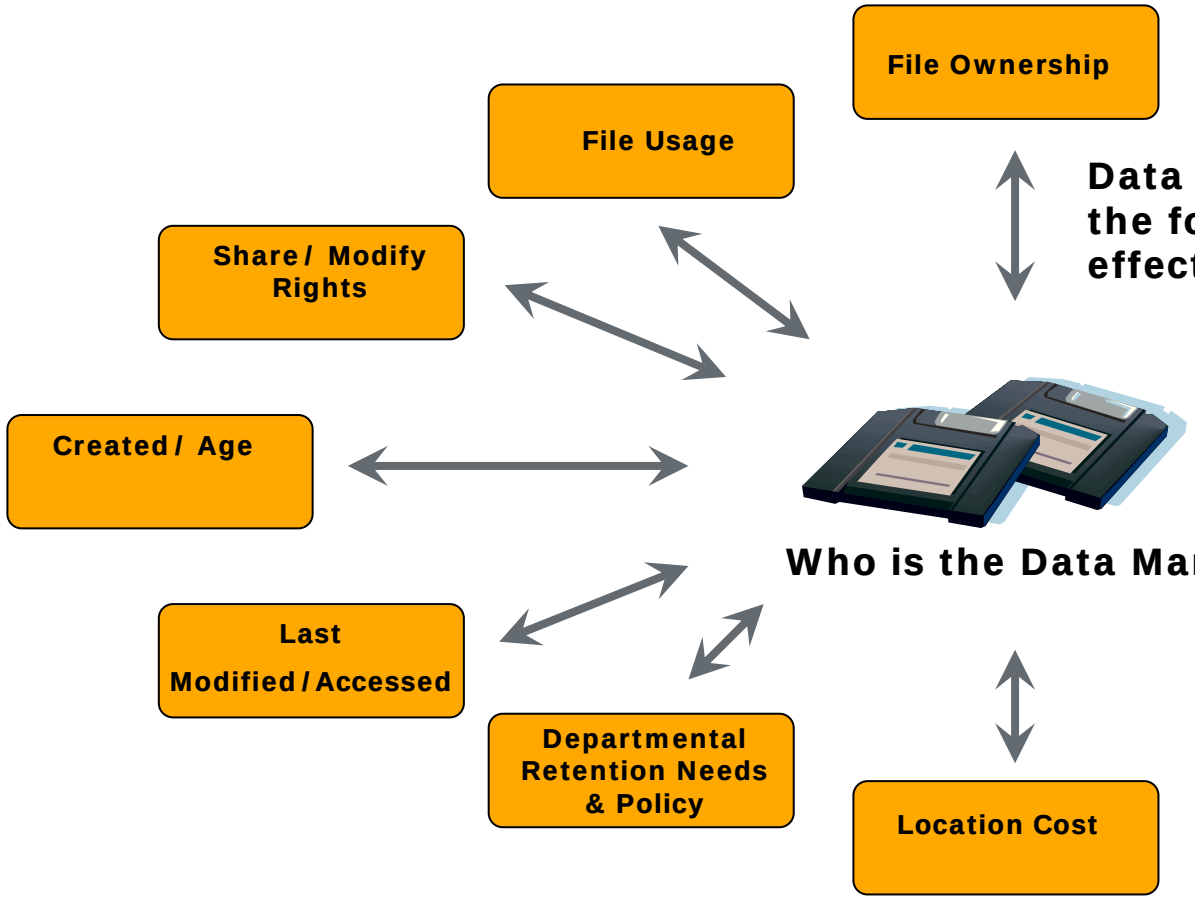
intermine FileCensus Instant Storage Insight

Who is managing your data growth?



Intermine expects Fortune 1000 companies to create a new IT position – the ILM Unstructured Data Manager

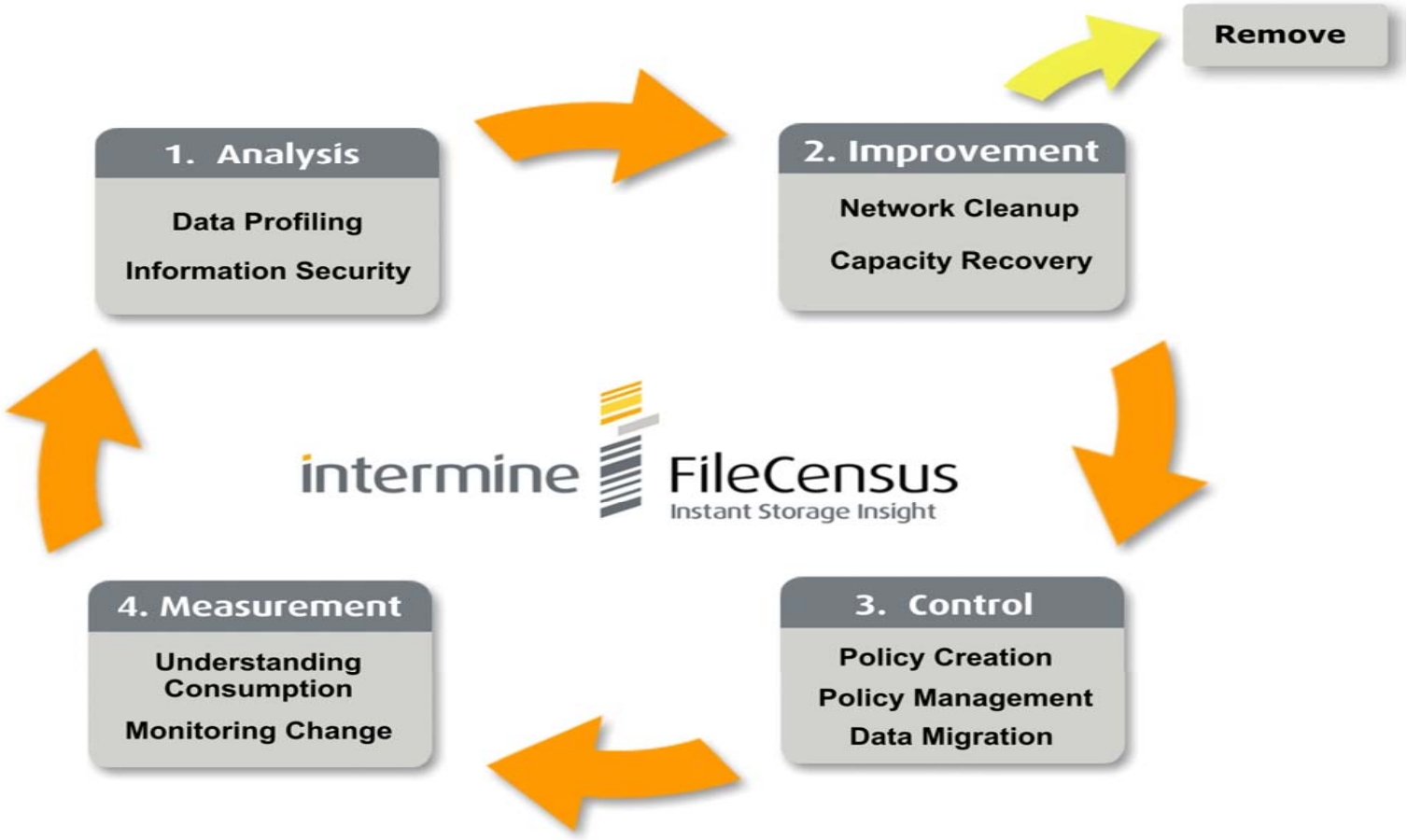
Managing Unstructured Data Needs



Data Profiling is the foundation for effective SRM

With Visibility - Assessment and Policy becomes possible – leading to automation.

Information is the real Business Asset



intermine FileCensus
Instant Storage Insight

FileCensus Overview

- ILM Solution – to manage Business Data not Storage Infrastructure while remaining complimentary to other SRM Infrastructure and HSM Software functionality
- FileCensus is the only product that works by scanning, compressing and keeping the details of every single file in the enterprise network daily and over time….
- Storage Insight into:
  - File Security
  - Redundant Data & Cleanup
  - Backup Insight & Improvement
  - Capacity Planning
  - Directory Base-lining
  - Departmental Chargeback (Access Based)
  - Automation – Data Migration & Control

FileCensus - How Does it Work?

- Heterogeneous - Agents scan Windows, NetWare and UNIX systems through the directory structure and recording details about every file and security information.
  - This meta-data is collected and stored daily.
- Unique IP - Compression technology allows for efficient collection and storage of this information to give unprecedented reporting speed and detail for daily or historical analysis.
  - ie: A Server OS (NT, Unix or NetWare) needs around 1,000 MB to manage 1,000,000 files. FileCensus equivalent data file is around 30 MB, allowing for daily snapshot capture/retention.
  - Competitors use traditional SQL Databases requiring around 500 MB per million files per day.
- Data Warehouse Approach to Analysis - Administrators and optionally Users can browse interactively and query the information, without placing loads on the production servers or guess what information is required beforehand.

FileCensus Architecture

Image Snapshots:

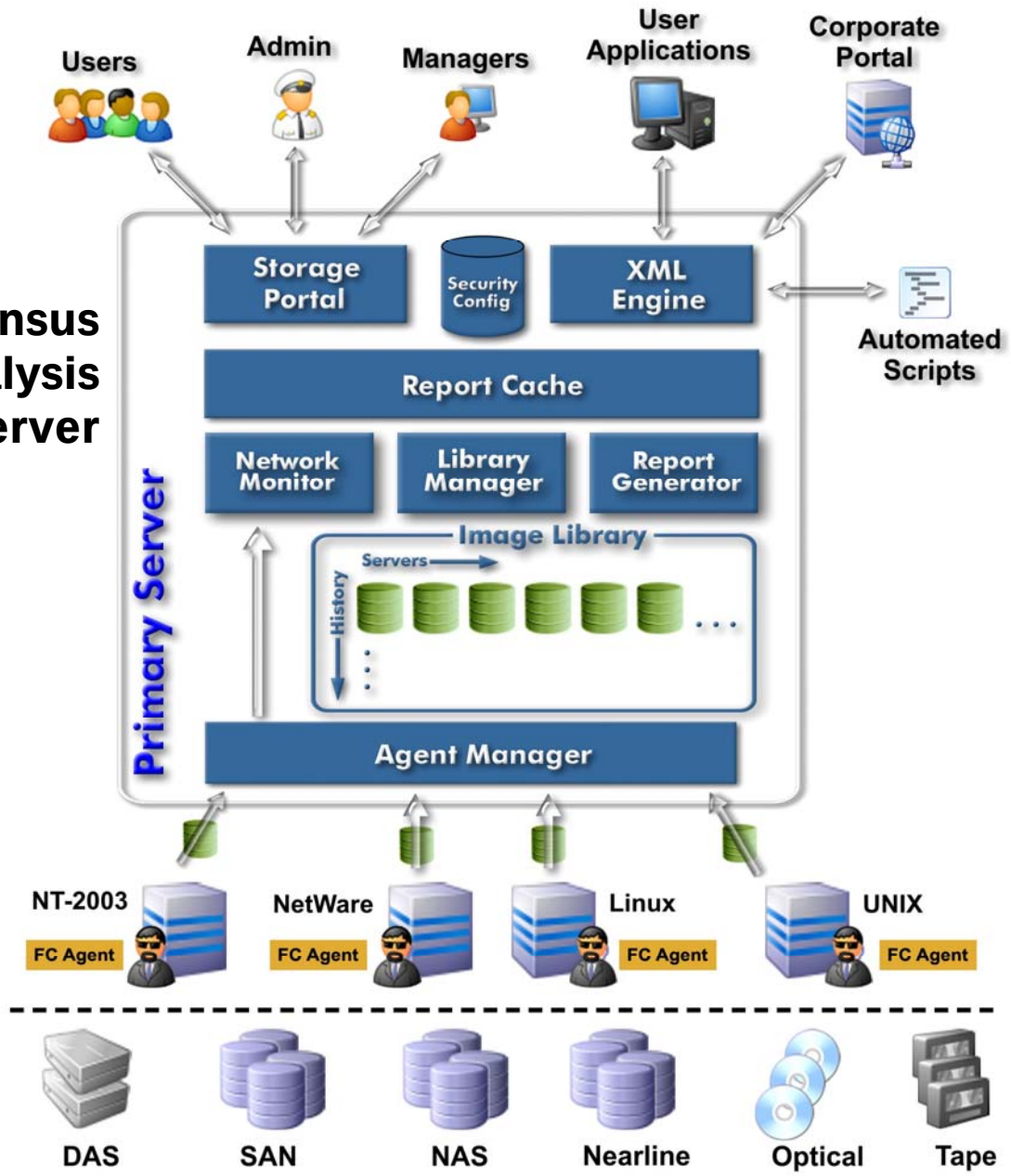
- Daily
- Weekly
- Monthly +

360 degree View

Data by:

- Age
- Type
- Size
- User
- Path
- Dept
- Classification
- Properties

FileCensus Analysis Server



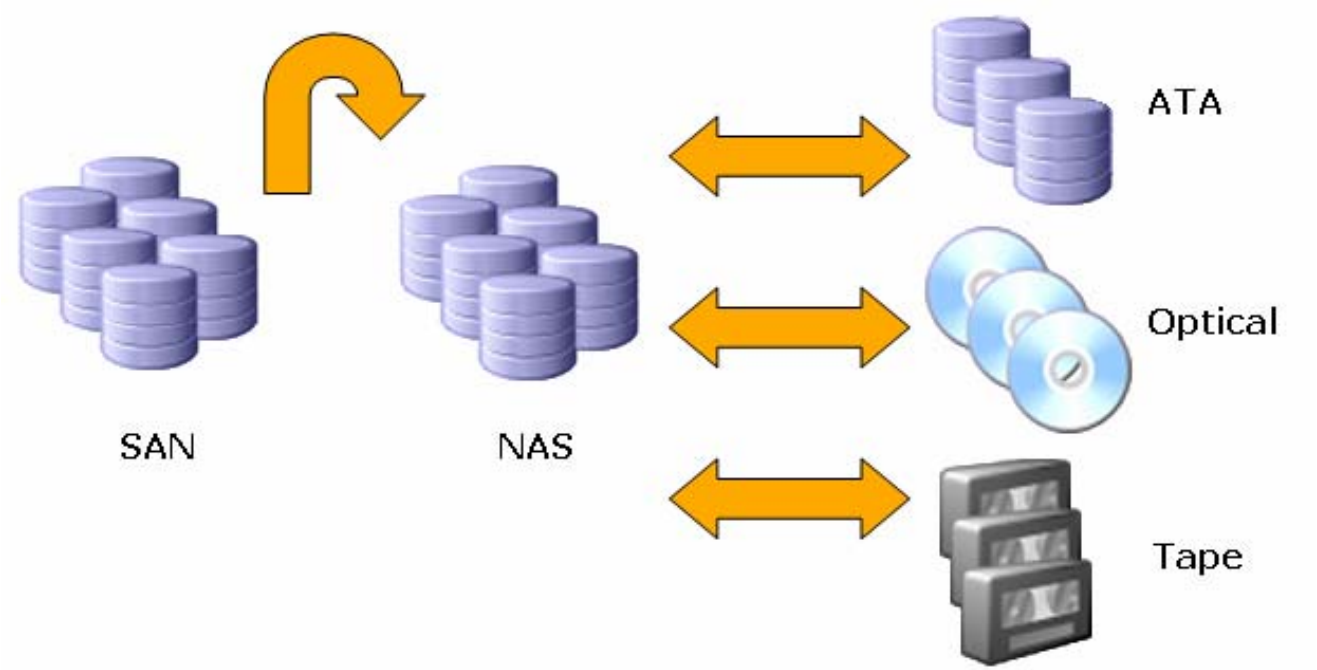
File Properties - Intermine Style

- Properties maintained by the storage systems:
  - Host, Path, Name, Size, Modified, Accessed, Created, Owner, Attributes
- Properties supplied by the applications and authors:
  - Title, Keywords, Author, Company, Category, User Defined Fields
- Properties calculated with daily FileCensus scans:
  - Access Patterns, Past Owners

intermine FileCensus
Instant Storage Insight

File Archival - Lowering Costs

- HSM - relocate data onto relevant storage platform according to value and policy. The benefits of moving data to cheaper storage platforms seems obvious.



- What data do you move to reduce TCO and how do you set policy?

File Archival – Who determines value?

	Data Type	Age	Access	Modify	Compliance
Legal	DOC	3 mths	✓		✓
Finance	XLS	6 mths		✓	✓
Marketing	PPT	3 mths	✓		
Corporate	PDF	1 mth	✓		✓
Funds Mgt	MDB	1 mth	✓		
↓	↓	↓	↓	↓	↓

- How many data types and lines of business do you have to consider before setting policy? How will you classify your data?

ROI - Return on Investment

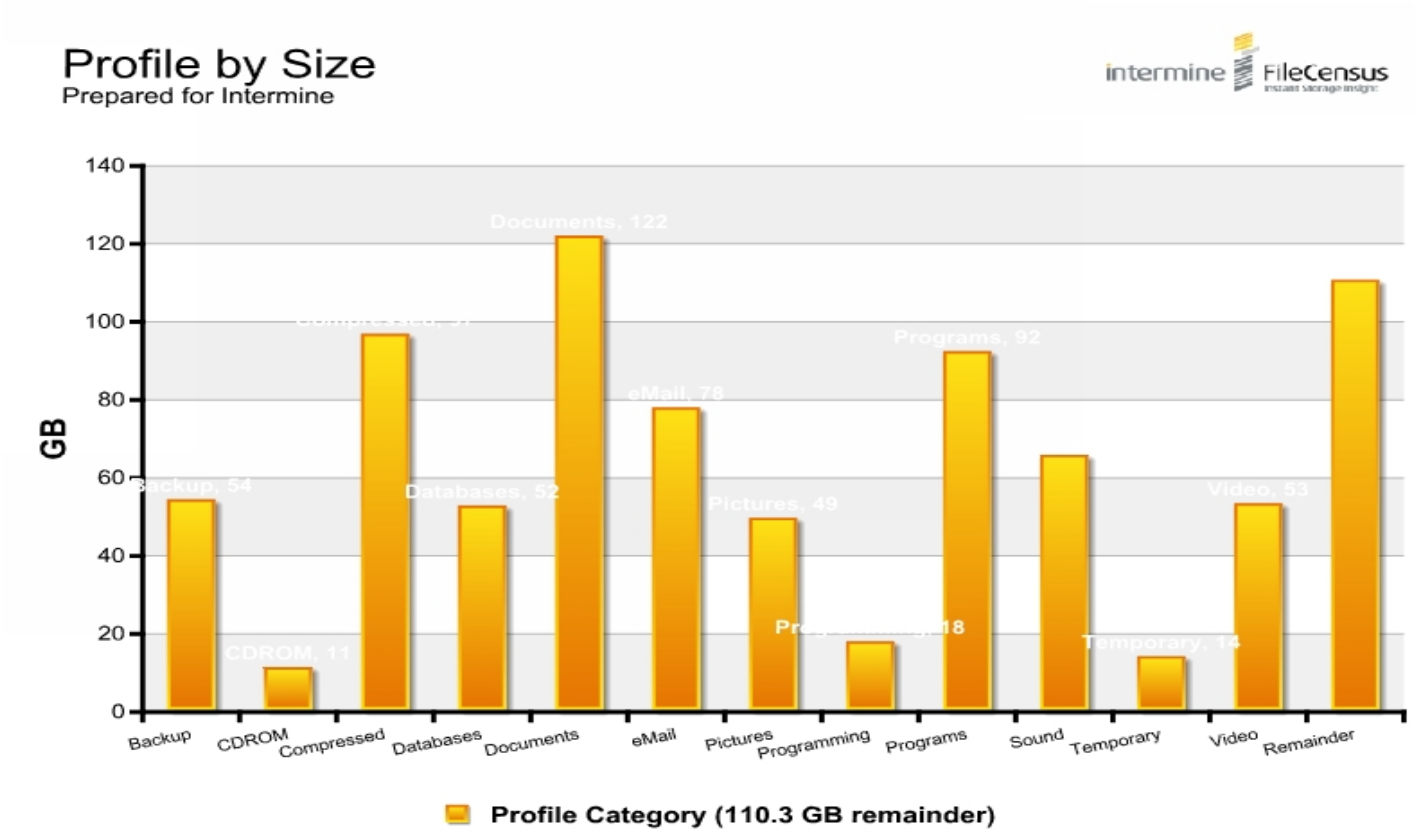
- FileCensus offers assistance with Data Discovery and Classification to help with planned migrations of unstructured data to appropriate storage platforms.

- By extrapolating simple metrics from the FileCensus data collected, we are able to demonstrate where to gain immediate and longer term ROI in the areas of:

  - Hardware Lifecycle Extension
  - Capacity Recovery
  - Data Migration
  - Improved Backup Management
  - Compliance Risk Mitigation
  - Storage TCO Reduction.

Example - Real US Customer

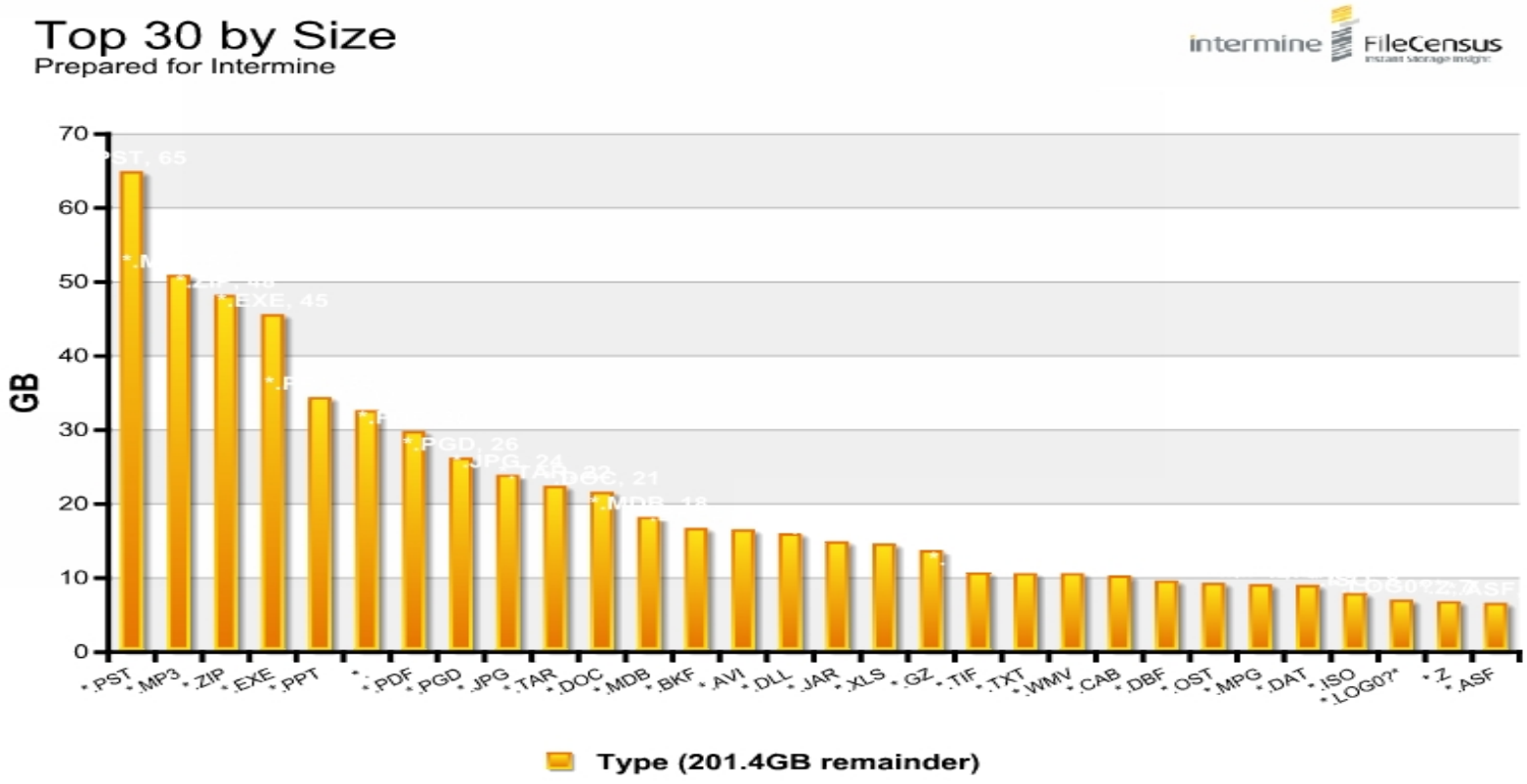
- Enterprise Customer – 500TB of data on SAN, NAS & DAS



intermine FileCensus
Instant Storage Insight

Example – by Application

- When customers first use FileCensus to analyze their unstructured data across an enterprise, a number of specific data patterns, growth trends and user behaviors become clearly apparent and usually revealed for the first time.

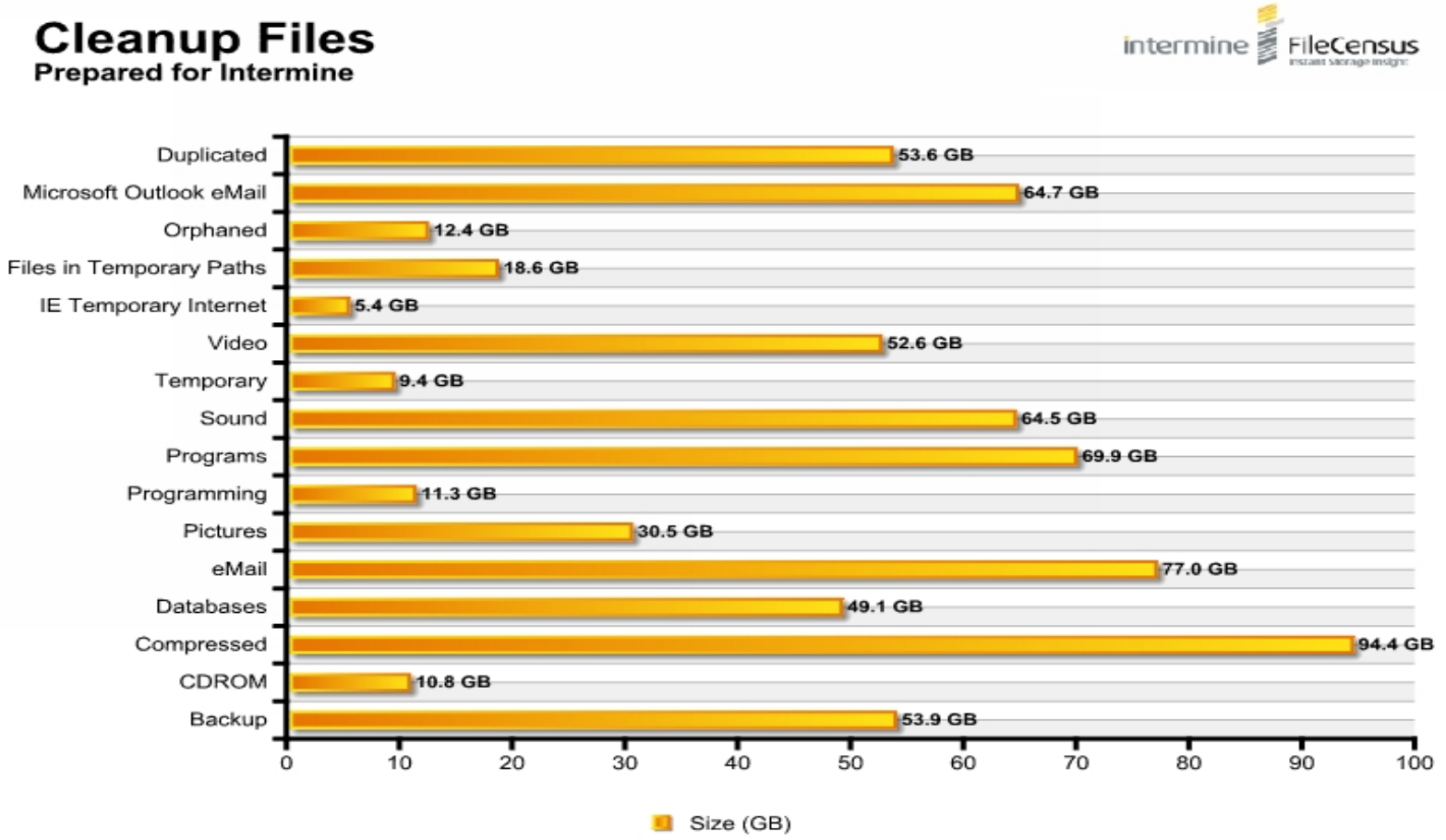


intermine FileCensus
Instant Storage Insight

Example – Regaining Capacity

- Each of the different data types and directories below can be examined in detail:
  - Recycle bins
  - Orphan files
  - Video, Sounds, Pictures, and Temp files
  - Duplicate files
  - Executables files in shared user areas
  - Runaway log files
  - Illegal and Inappropriate Files

Example – Regaining Capacity

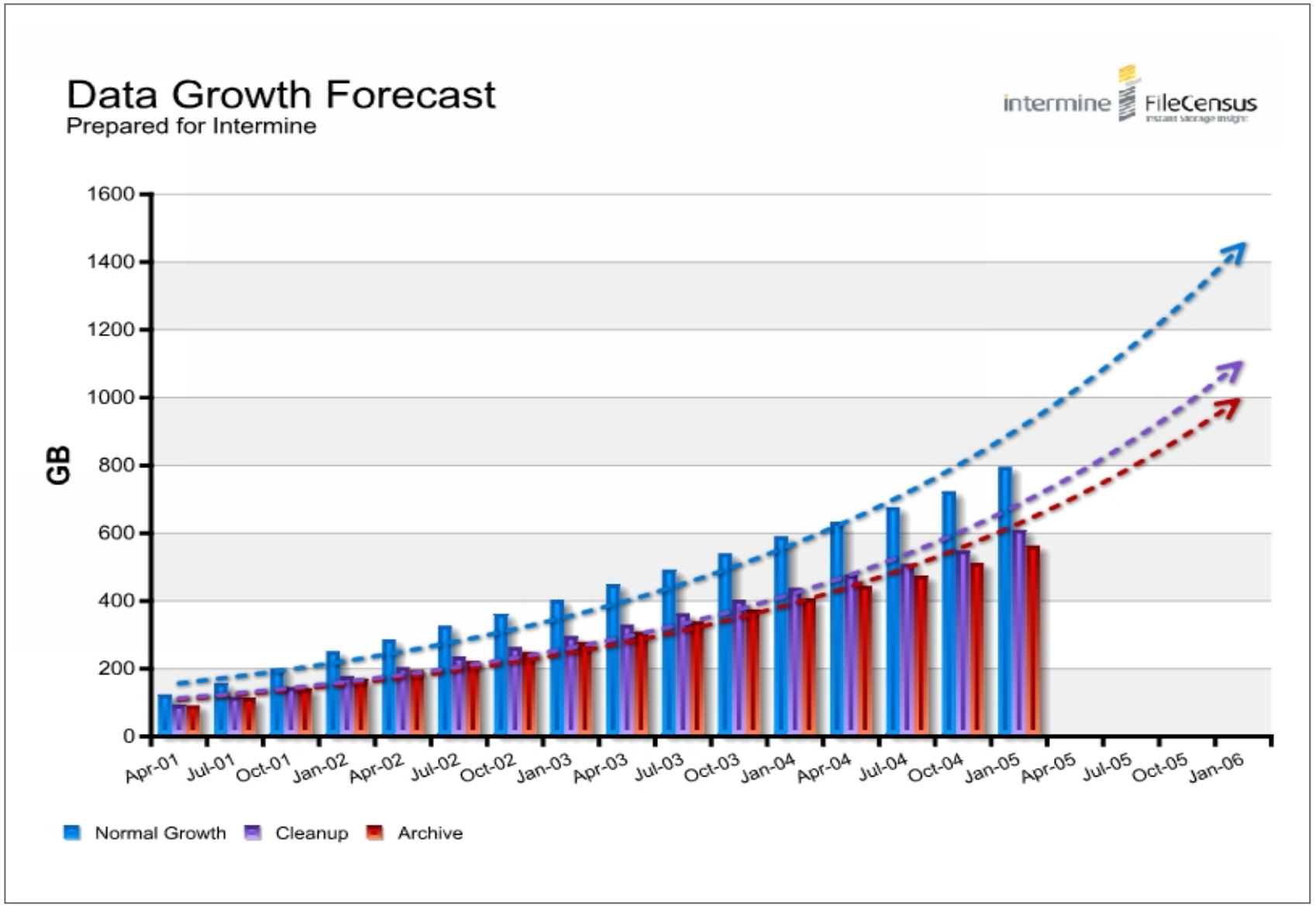


intermine FileCensus
Instant Storage Insight

Example – Where to From Here?

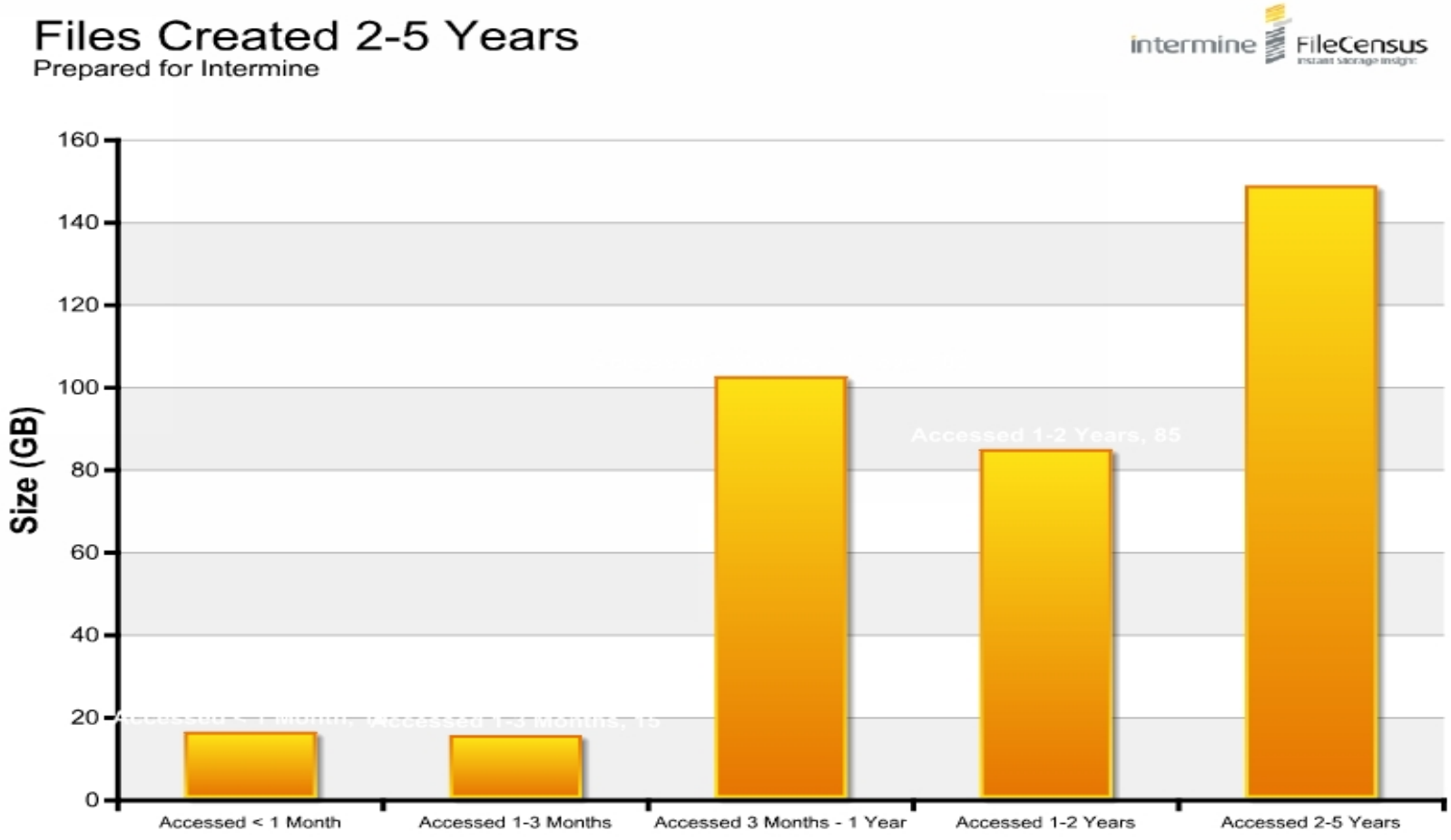
- Once the fire fighting is under control (Security, Cleanup and Backup) you can turn to understanding how the storage is really being used. Many of the reports can be used to highlight populations of files hidden deep within directory structures or spread across many volumes.

- Before you start capacity planning and deciding which data needs to be moved off to tape, compliance reference systems or cheaper disk to reduce overall TCO, you need to examine the quality of your meta-data concerning true access dates, access frequency and file ownership.



intermine FileCensus
Instant Storage Insight

Example – Data Usage Trends



intermine FileCensus
Instant Storage Insight

Example – Migration Findings

- Average Growth Rate – 50% Year
  - 2000 to 2001 = 100%
  - 2001 to 2002 = 20%
  - 2002 to 2003 = 35%
  - 2003 to 2004 = 70% (so far)
- Minor Access Date corruption – 3 years ago
- 1 Year Since Accessed = 34% of Data
- Approx 32% of data created more than 2 years ago.
- Capacity Limit – Feb 2005 based on current growth
- Need to classify data move to appropriate tiers to reduce TCO
- Corporate Policy and Control will reduce Storage TCO.

Example – Improvement & Data Migration

- Current Primary Tier Capacity - 500 TB
  - 500TB @ $ 50,000 per TB Average
  - $700K / month on 3 year lease
  - Storage services = $2.1M month in managed storage services
  - Savings (Capacity) after Improvement = 33% = $924,000 per month
- Migration Modeling

	Current	Future
Primary Storage ($50GB)	300TB	200TB
Secondary ($40GB)	200TB	100TB
Nearline ($20GB)	0TB	200TB

  - Hardware Savings (Migration) after improvement = 23% = $161K /mth lease.
  - Storage Service Savings (Migration) after improvement = 23% = $483K/ mth
- Total Savings – after improvement - $1.5 Million per month

FileCensus - ROI

The FileCensus return on investment can be calculated by:

- Immediate Storage Capacity Recovery through Cleanup
- Reduced Costs by introducing Data Migration to cheaper storage tiers
- New internal Chargeback System
- Impact of chaos, compliance and growth without Visibility and Control

While always delivering - Rapid Time to Value

intermine FileCensus
Instant Storage Insight

Thank you
for your time