IBM INFORMATION INTEGRATION & GOVERNANCE SYMPOSIUM 2012

Delivering Trusted Information for Smarter Business Decisions

Control your data, Don't let it control you!

Ben Davis- Information Governance Specialist

1 May 2012

Our World today

or mention #IIGS

Big Data!

Some facts

- In 2009, amid the "Great Recession," the amount of digital information grew 62% over 2008 to 800 billion gigabytes (0.8 Zettabytes).

- 75 billion fully-loaded 16 GB Apple iPads, which would fill the entire area of Wembley Stadium to the brim 41 times

- The amount of digital information created annually will grow by a factor of 44 from 2009 to 2020, as all major forms of media - voice, TV, radio, print - complete the journey from analog to digital.

- By 2020, the percent of digital information requiring security beyond baseline levels will grow from 30% to 50%.

- 35% more digital information is created today than the capacity exists to store it. This number will jump to over 60% over the next several years.

Source: **The Digital Universe Decade - Are You Ready?- IDC May 2010**

twitter: *Follow @ANZ_IM or mention #IIGS*

Organizations have been increasingly challenged with successfully managing data growth

IBM®

Increasing Costs

3-10x

Cost of managing storage over the cost to procure[a]

$1.1 billion

Amount organizations will have spent in 2011 on storage[b]

Poor Application Performance

80%

The time DBA's spend weekly on disk capacity issues[c]

250 hours

The amount of time needed to run "daily" batch processes[d]

Manage Risk & Compliance

50%

of firms retain structured data for 7+ years[e]

57%

of firms use Back-up for data retention needs[e]

(a) Merv Adrian, IT Market Strategies, "Data Growth Challenges Demand Proactive Data Management", November 2009
(b) IDC, "Worldwide Archival Storage Solutions 2011–2015 Forecast: Archiving Needs Thrive in an Information-Thirsty World", October 2011
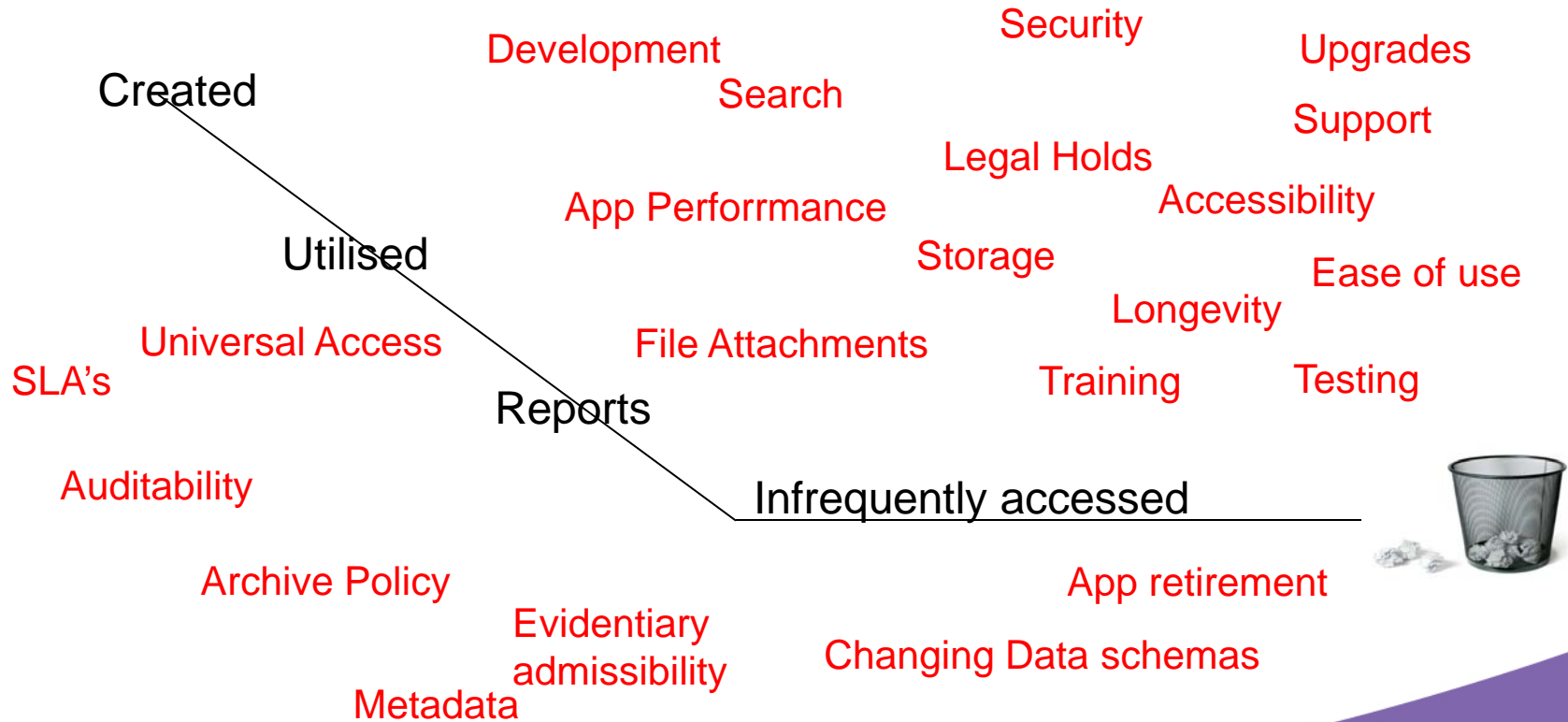(c) Simple-Talk, "Managing Data Growth in SQL Server ", January 2010
(d) IBM Client Case Study: *Toshiba TEC Europe*; archiving reduced batch process time by 75%
(e) IDC Quick Poll Survey 2011, "Data Management for IT Optimization and Compliance", November 2011

twitter: *Follow @ANZ_IM or mention #IIGS*

The Lifecycle of data

Created

Utilised

Reports

Infrequently accessed

Development

Security

Upgrades

Search

Support

Legal Holds

App Performrmance

Accessibility

Storage

Ease of use

Longevity

Universal Access

File Attachments

Training

Testing

SLA's

Auditability

Archive Policy

App retirement

Evidentiary admissibility

Changing Data schemas

Metadata

So when will this data storm end?

- The answer is never
    - Data will get bigger
    - Data will become more complex
- You cannot build a dam big enough to hold all of your data
- In the future you will be required to protect all of the data if not already!

So let's control our data, not let it control us

twitter: *Follow @ANZ_IM or mention #IIGS*

Control Data throughout the data lifecycle

| Information Governance Core Disciplines Lifecycle Management			
Discover & Define	**Develop & Test**	**Optimize & Archive**	**Consolidate & Retire**
Discovering and understanding the "what & where" of enterprise data	Developing models and code to store and access enterprise data including the configuration of data for test environments	Optimizing performance through identification of bottlenecks and building the right strategy for managing applications and data growth	Implementing a consistent process for retiring or consolidating applications

9

You can't govern what you don't understand
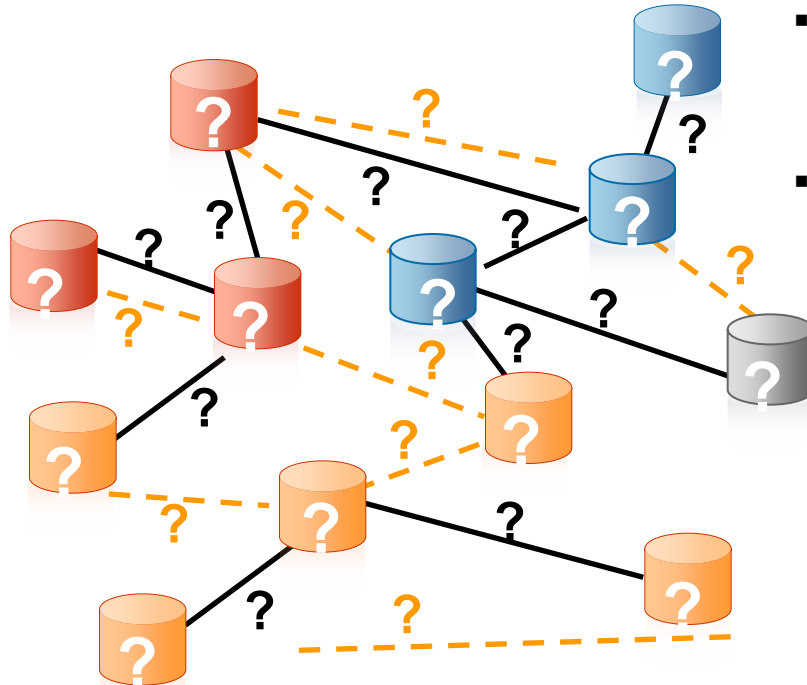
Distributed Data Landscape

- Data can be distributed over multiple applications, databases and platforms
  - Where are those databases located?
- Complex, poorly documented data relationships
  - Which data is sensitive, and which can be shared?
  - Whole and partial sensitive data elements can be found in hundreds of tables and fields

Data relationships not understood because:
  - Corporate memory is poor
  - Documentation is poor or nonexistent
  - Logical relationships (enforced through application logic or business rules) are hidden

twitter: *Follow @ANZ_IM or mention #IIGS*

Organizations continue to be challenged with building quality applications
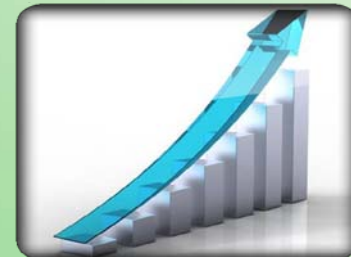
Increasing Risk

Mandatory to protect data and comply with regulations

Time to Market

Lack of realistic test data and inadequate environments

Increasing Costs

Defects are caught late in the cycle

11

Organizations continue to be challenged with building quality applications

Increasing Risk

45,000+
Number of sensitive records exposed to 3rd party during testing[c]

62%
companies use actual customer data to test applications[a]

Time to Market

37%
Satisfied with speed of software development[f]

30-50%
Time testing teams spend on setting up test environments, instead of testing[b]

Increasing Costs

$300 billion
Annual costs of software-related downtime.[d]

32%
Low success rate for software projects[e]

a. The Ponemon Institute. The Insecurity of Test Data: The Unseen Crisis
b. NIST, Planning Report. The Economic Impacts of Inadequate Infrastructure for Software Testing
c. Federal Aviation Administration: Exposes unprotected test data to a third party http://fcw.com/articles/2009/02/10/faa-data-breach.aspx
d. The Standish Group, *Comparative Economic Normalization Technology Study,* CHAOS Chronicles v12.3.9, June 30, 2008
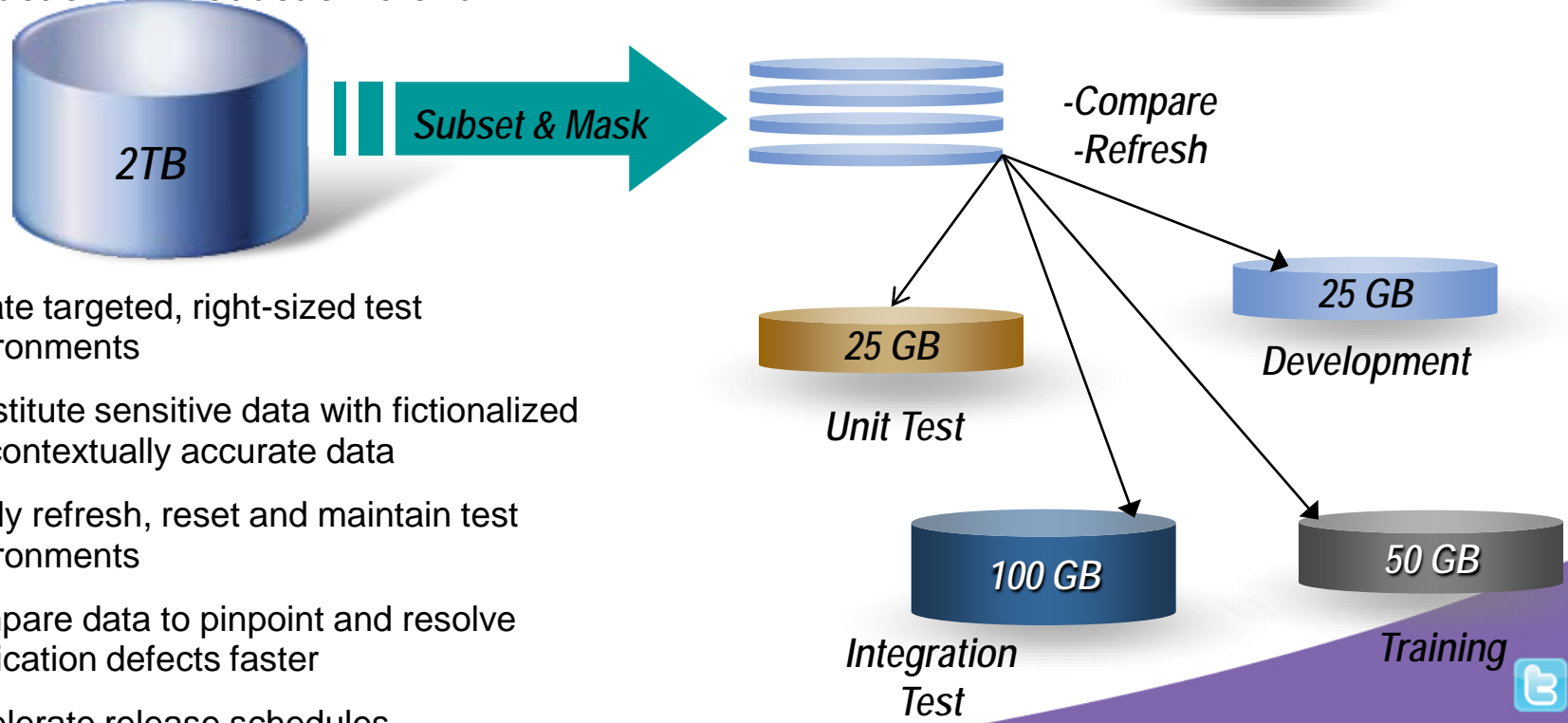e. The Standish Group, Chaos Report, April 2009
f. Forrester Research, *"Corporate Software Development Fails To Satisfy On Speed Or Quality",* 2005

12

Employ effective test data management practices

Information Governance Core Disciplines
Lifecycle Management

Disco Def	**Develop & Test**	te & e	Consolidate & Retire

Production or Production Clone

2TB

Subset & Mask

-Compare
-Refresh

25 GB

Development

25 GB

Unit Test

100 GB

Integration Test

50 GB

Training

- Create targeted, right-sized test environments

- Substitute sensitive data with fictionalized yet contextually accurate data

- Easily refresh, reset and maintain test environments

- Compare data to pinpoint and resolve application defects faster
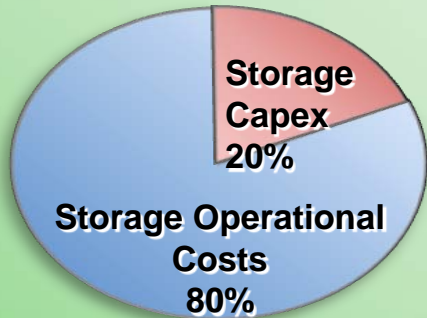
- Accelerate release schedules

13

Improve application quality and delivery efficiency

Reduce Cost	Reduce Risk	Speed Delivery
Automate creation of realistic "right sized" test data to reduce the size of test environments	**Mask** sensitive information for compliance and protection	**Refresh** test data through self-service, speeding testing and application delivery

- Understand what test data is needed for test cases
- Create "right-sized" test data by subsetting
- Ensure masked data is contextually appropriate to the data it replaced, so as not to impede testing
- Easily refresh & maintain test environments through self service access by developers and testers
- Automate test result comparisons to identify hidden errors
- Support for custom & packaged ERP applications in heterogeneous environments

twitter: *Follow @ANZ_IM or mention #IIGS*

Organizations have been increasingly challenged with successfully managing data growth

IBM

Information Governance Core Disciplines
Lifecycle Management

| Discover& Define | Deve T | **Optimize & Archive** | date & re |

Increasing Costs

Storage Capex 20%

Storage Operational Costs 80%

Buying more storage is not a "cheap" fix when you add the operational burden

Poor Application Performance

Application Performance
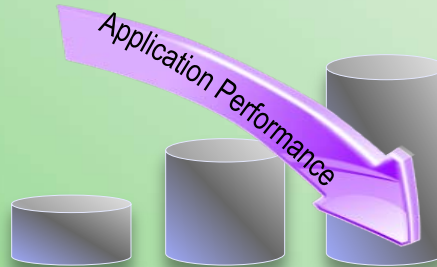
Business users & customers wait for application response; DBA's spend majority of time fixing performance issues

Manage Risk & Compliance

The "keep everything" strategy can impact disaster recovery and data retention & disposal compliance

twitter: *Follow @ANZ_IM or mention #IIGS*

Organizations have been increasingly challenged with successfully managing data growth

IBM

Information Governance Core Disciplines
Lifecycle Management

| Discover & Define | Dev... T... | **Optimize & Archive** | ...date & ...re |

Increasing Costs

3-10x
Cost of managing storage over the cost to procure[a]

$1.1 billion
Amount organizations will have spent in 2011 on storage[b]

Poor Application Performance

80%
The time DBA's spend weekly on disk capacity issues[c]

250 hours
The amount of time needed to run "daily" batch processes[d]

Manage Risk & Compliance

50%
of firms retain structured data for 7+ years[e]

57%
of firms use Back-up for data retention needs[e]

(a) Merv Adrian, IT Market Strategies, "Data Growth Challenges Demand Proactive Data Management", November 2009
(b) IDC, "Worldwide Archival Storage Solutions 2011–2015 Forecast: Archiving Needs Thrive in an Information-Thirsty World", October 2011
(c) Simple-Talk, "Managing Data Growth in SQL Server", January 2010
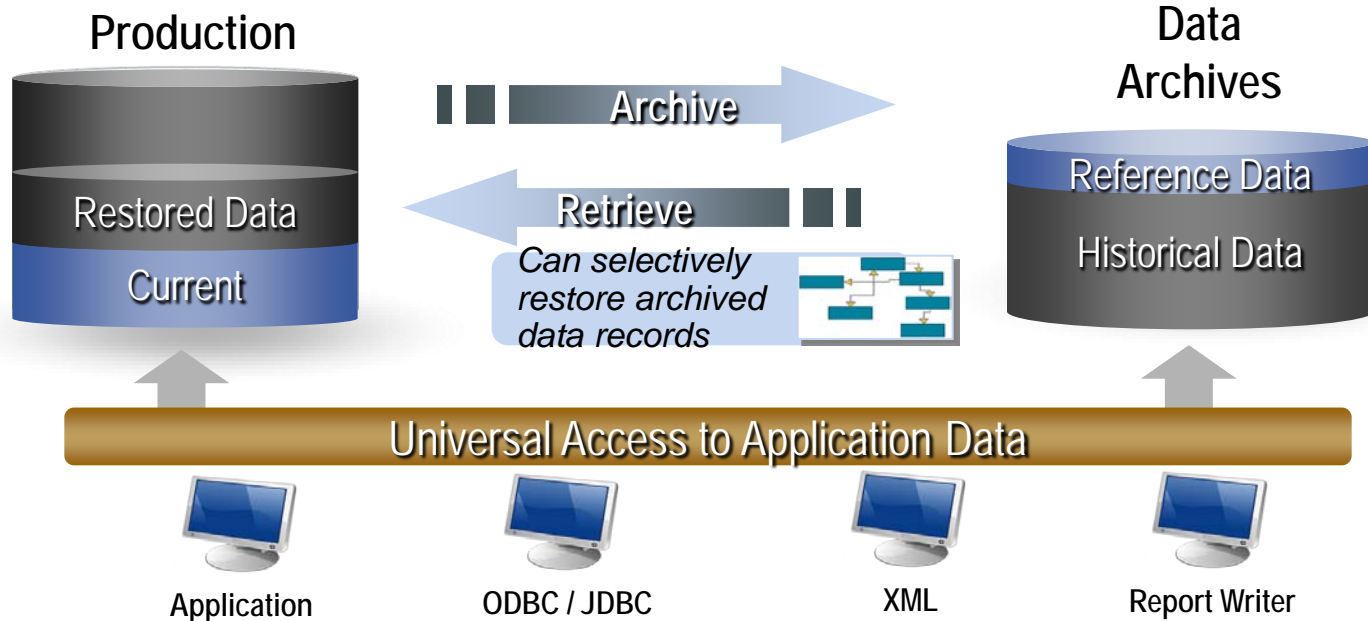(d) IBM Client Case Study: Toshiba TEC Europe; archiving reduced batch process time by 75%
(e) IDC Quick Poll Survey 2011, "Data Management for IT Optimization and Compliance", November 2011

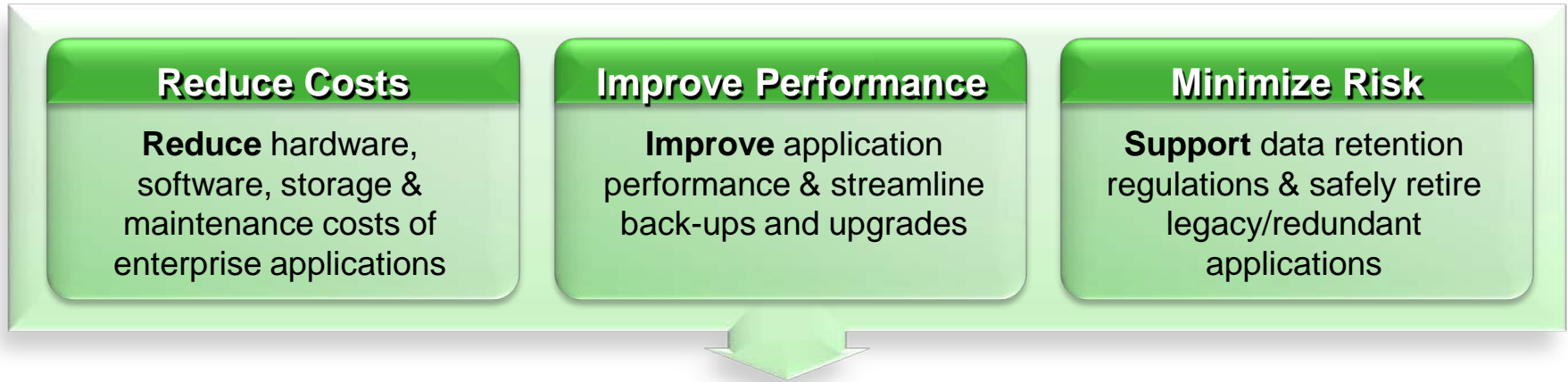twitter: *Follow @ANZ_IM or mention #IIGS*

Archive historical data for data growth management

Information Governance Core Disciplines
Lifecycle Management

Discover&
Define

Dev...
T...

**Optimize &
Archive**

date &
re

IBM.

Production

Data
Archives

Archive

Restored Data

Retrieve

Current

*Can selectively
restore archived
data records*

Reference Data

Historical Data

Universal Access to Application Data

Application

ODBC / JDBC

XML

Report Writer

17

Effectively Archive and Manage Data Growth

IBM

| Reduce Costs	Improve Performance	Minimize Risk
Reduce hardware, software, storage & maintenance costs of enterprise applications	**Improve** application performance & streamline back-ups and upgrades	**Support** data retention regulations & safely retire legacy/redundant applications

- Discover & identify data record types to archive across heterogeneous environments
- Intelligently archive data to improve application performance and support data retention
- Capture & store historical data in its original business context
- Define & maintain data retention policies consistently across the enterprise
- Ensure long-term, application-independent access of archived data via multiple access methods
- Support for custom & packaged ERP applications in heterogeneous environments

When it's time to retire or consolidate applications

- Application portfolio has redundant systems acquired via mergers and acquisitions
- Line of business divested; application is no longer needed
- Legacy technologies not compatible with current IT direction
  - Old database and/or application versions no longer supported by manufacturer
- Required technical skills or application knowledge no longer available
- Budget pressures – do more with less
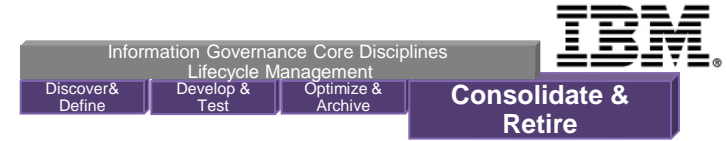
In almost ALL cases, access to legacy data MUST be retained while the application and database are eliminated

twitter: *Follow @ANZ_IM or mention #IIGS*

Retire redundant and legacy applications

- Preserve application data in its business context
  - Capture all related data, including transaction details, reference data & associated metadata
  - Capture any related reference data that may reside in other application databases
- Retire out-of-date packaged applications as well as legacy custom applications
  - Leverage out-of-box support of packaged applications to quickly identify & extract the complete business object
- Shut down legacy system without a replacement
  - Provide fast and easy retrieval of data for research and reporting, as well as audits and e-discovery requests

Infrastructure before Retirement

User — Application — Database — Data

User — Application — Database — Data

User — Application — Database — Data

Archived Data after Consolidation

User, User, User — Archive Engine — Archive Data

twitter: *Follow @ANZ_IM or mention #IIGS*
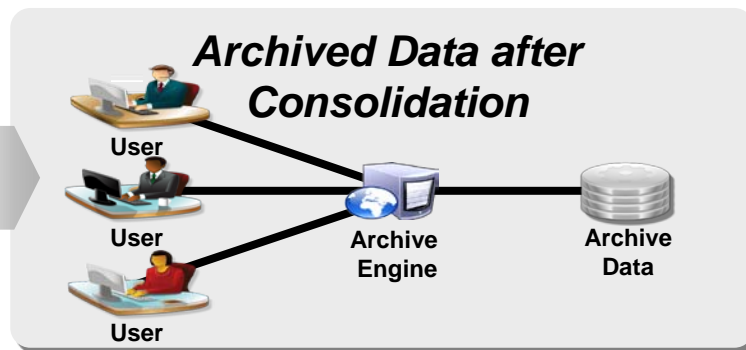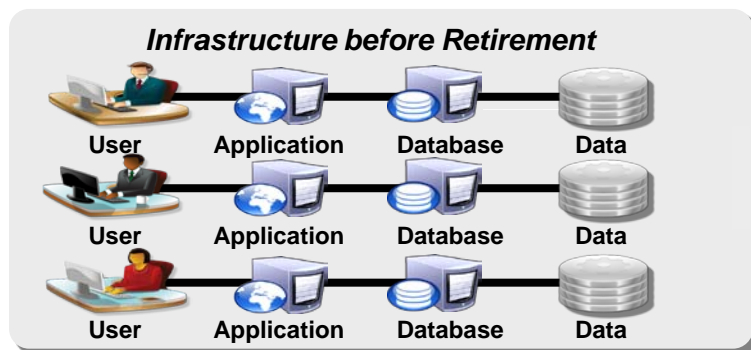
IBM meeting these challenges

Non Production Environments
Subset & Mask

Siebel | Oracle | PeopleSoft | JDEdwards | SAP | Amdocs | Custom | OEM/ISV

Production Environments
Archive

Siebel | Oracle | PeopleSoft | JDEdwards | SAP | Amdocs | Custom | OEM/ISV

Data Growth, Data Privacy, Test Data Management, Application Upgrades, Application Retirement

Optim ™

| Oracle | SQL Server | Sybase | Informix | DB2 LUW | XML | IMS | VSAM | Adabas | DB2 z/OS | Teradata |

Windows XP/2000   Solaris   HP/UX   Linux   AIX   OS/390   Z/OS   i-Series

NAS   SAN   ATA   CAS   Optical   Tape

Single, scalable, interoperable EDM solution provides a central point to deploy policies to extract, store, port, and protect application data records from creation to deletion

Questions?

<u>Further Information</u>

www.optimsolution.com

Ben Davis

Benjdav@au1.ibm.com

twitter: *Follow @ANZ_IM or mention #IIGS*