Removing Data Constraints in Testing

Todd Blatti
Director, Solution Engineering

June 2017

Common Trade-offs / Criteria in Test Data Management

1 Speed: increasing testing agility

2 Quality: improving testing coverage

3 Security: ensuring data privacy

4 Cost: optimizing infrastructure resources

Collect

Apps

Files

Control

Databases

Consume

Drivers for Test Data Management

20%
of a tester's time is spent on test data-related activities.

- Cognizant 20-20 Insights

TEST DATA CHALLENGES

33%
of application projects exceed their schedules.

- McKinsey Survey

LACK OF **APPLICATION** SPEED & AGILITY

$3.8M
is the average cost of a data breach across industries.

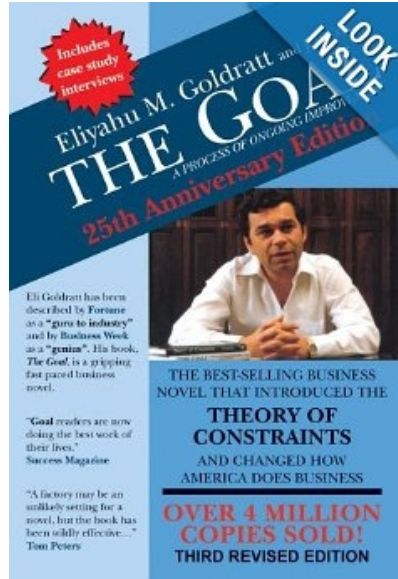- Ponemon Institute Report

IMPACT OF **BUSINESS** FAILURES

Evaluating Common Test Data Approaches

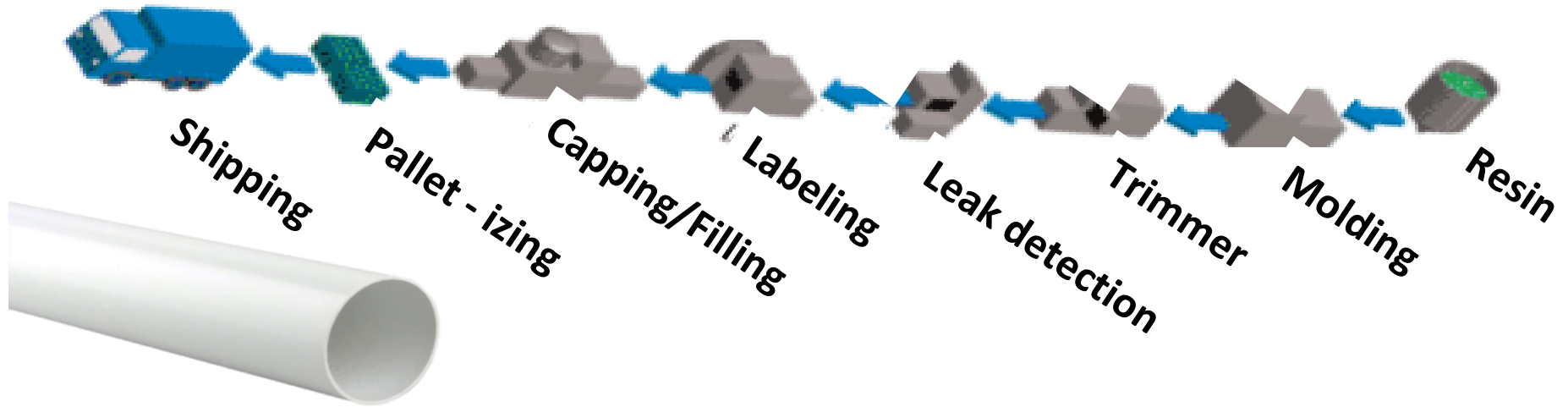| Test Data Type vs. Criteria	Speed	Quality	Security	Infrastructure Costs
Production Data	Slow, complex, manual processes	Improved test coverage and representativeness	Sensitive data at risk in non-production	High consumption of storage, CPU, and licenses
Subset of Production	Less time to provision than full production copies	Missed test case outliers	Sensitive data at risk in non-production	Some hardware, CPU, and license savings
Masked Data (Full or Subset)	Extended SLAs for masked data	Must maintain referential integrity	Improved data security	Requires standalone masking solution and staging server
Synthetic Data	Risk of human error creating test data	Limited to a small percentage of testing requirements	No requirement for de-identification of sensitive data	Limited space savings

Red = Fails to meet criteria   Yellow = Partially meets criteria   Green = Meets criteria
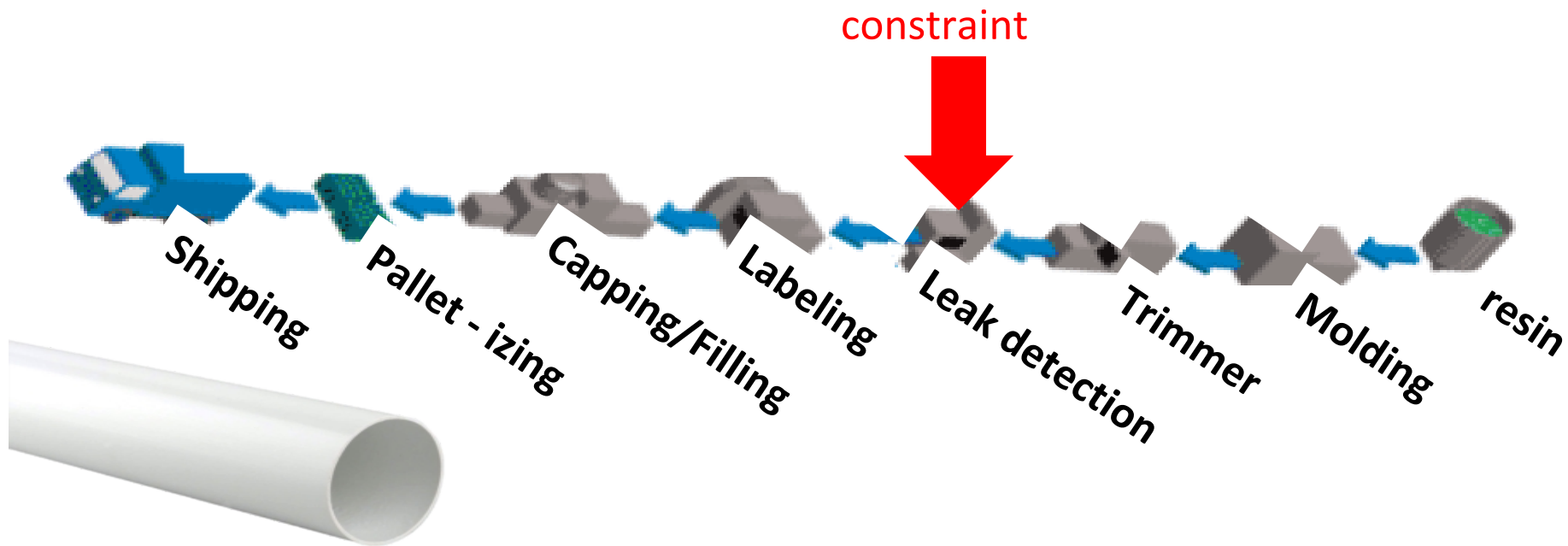
4

The Goal: Theory of Constraints

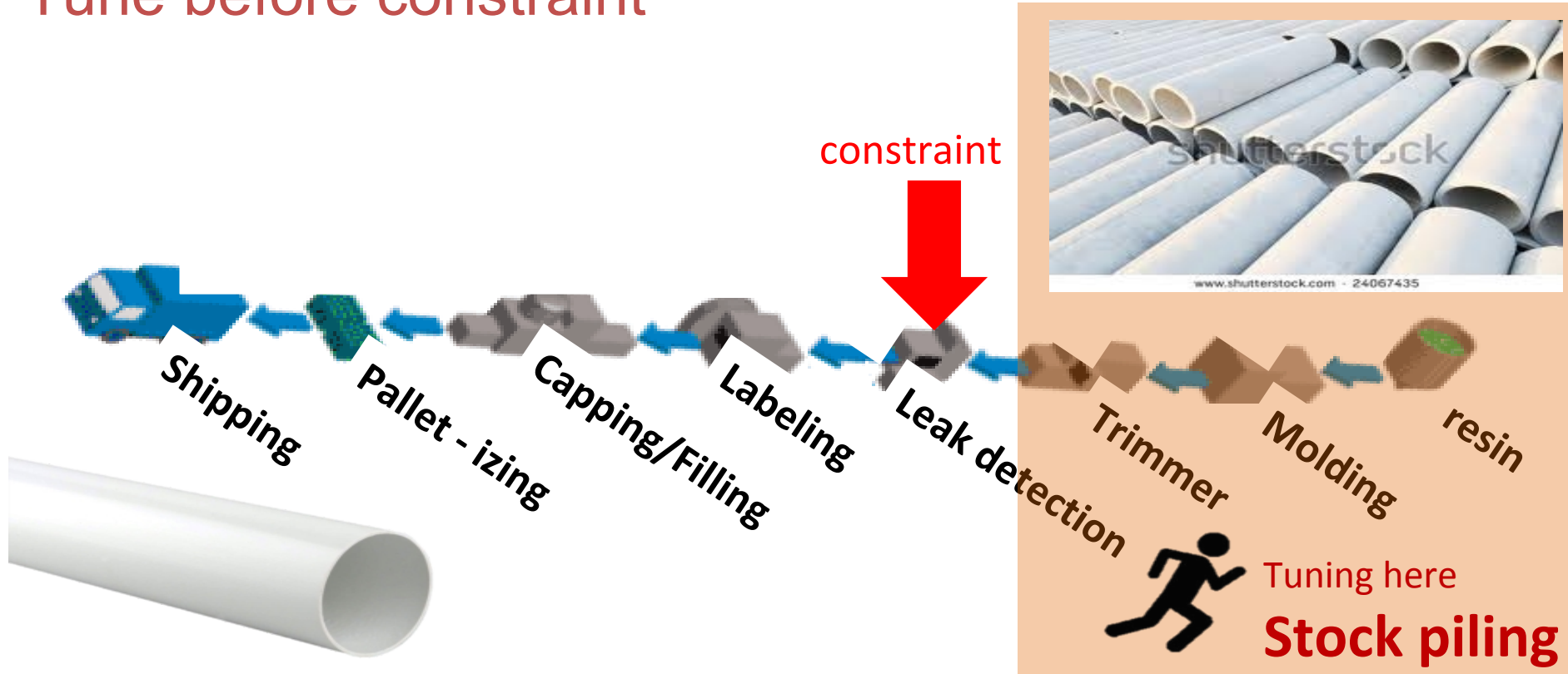- **Improvement** not made at the **constraint** is an **illusion**
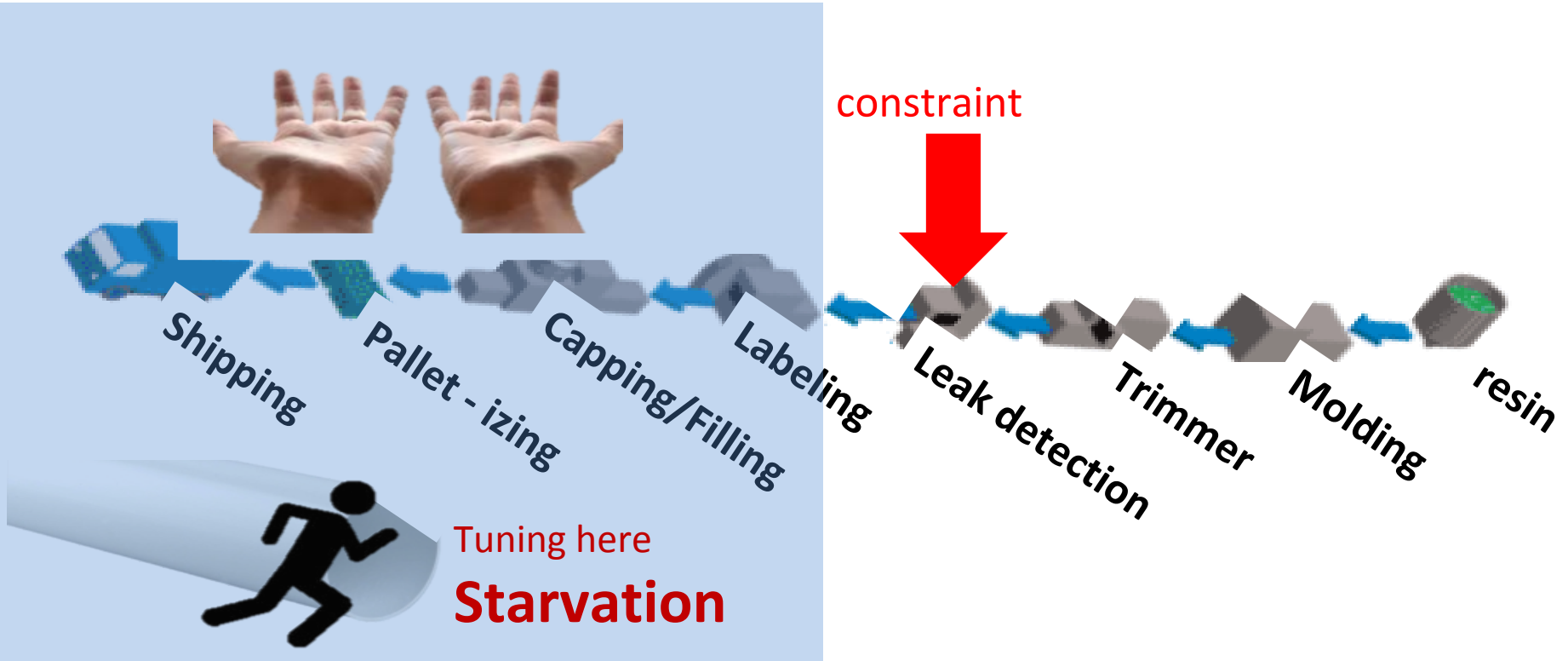
Factory floor

Shipping Pallet - izing Capping/Filling Labeling Leak detection Trimmer Molding Resin

Factory floor

constraint

Shipping · Pallet - izing · Capping/Filling · Labeling · Leak detection · Trimmer · Molding · resin

Tune before constraint

constraint

Shipping

Pallet - izing

Capping/Filling

Labeling

Leak detection

Trimmer

Molding

resin

Tuning here

Stock piling

Tune after constraint

constraint

Shipping

Pallet - izing

Capping/Filling

Labeling

Leak detection

Trimmer

Molding

resin

Tuning here
Starvation

Factory floor: optimizing at the constraint

constraint
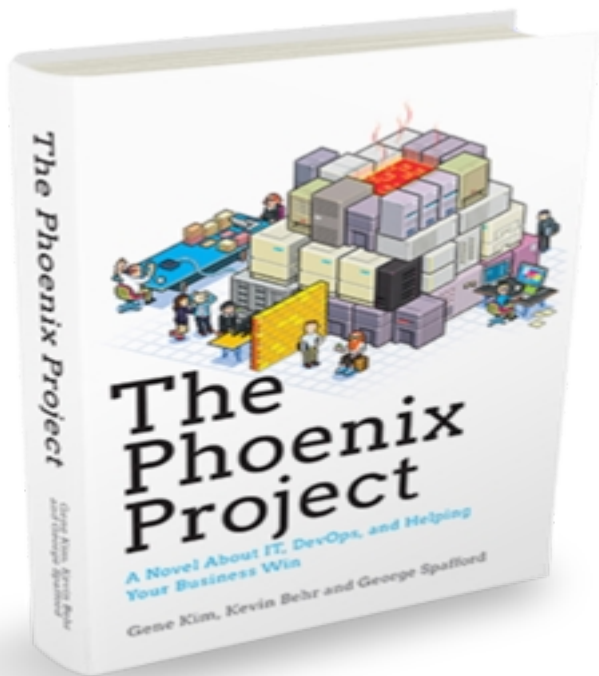
Shipping Pallet - izing Capping/Filling Labeling Leak detection Trimmer Molding resin
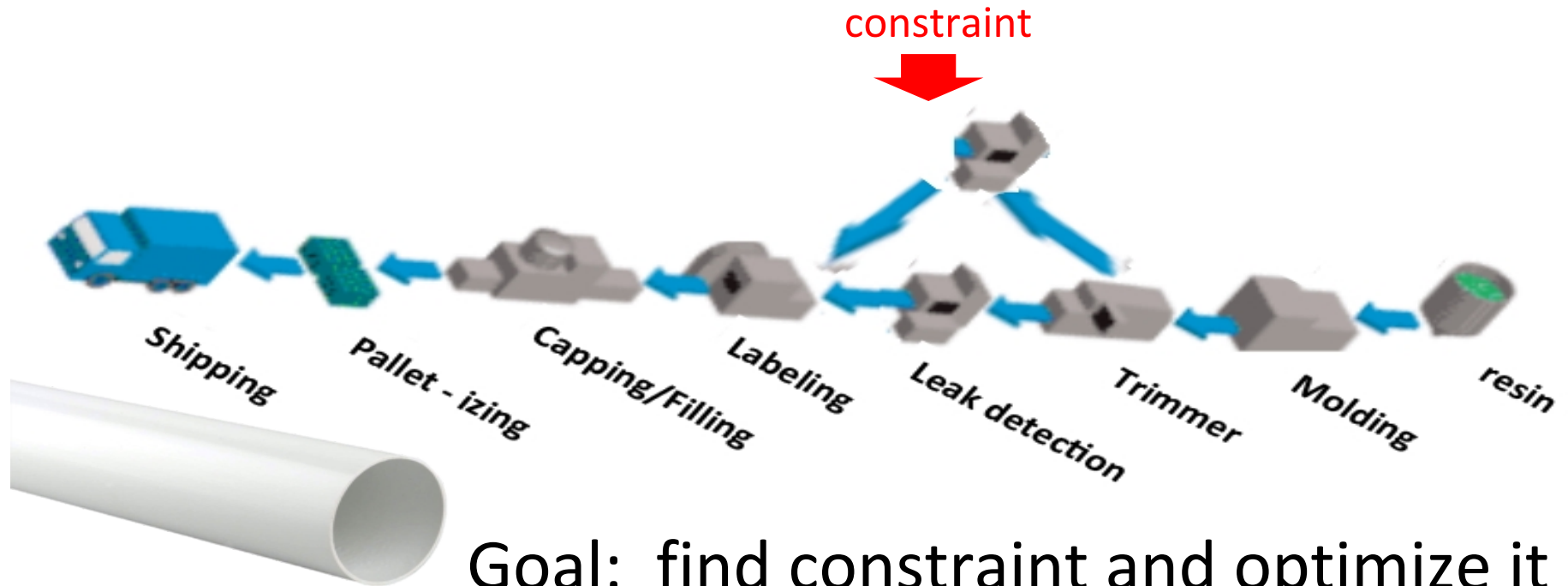
Goal: find constraint and optimize it

Does the Theory of Constraints work for IT ?

- Goals Clarify
- Metrics Define
- Constraints Identify
- Priorities Set
- Iterations Fast

- Continuous Integration
- Cloud
- Agile
- Kanban
- Kata

"IT is the factory floor of this century"

The Phoenix Project: The Goal applied to IT

What is the constraint in IT ?

"IT is the factory floor of this century"

The Phoenix Project: top 5 constraints in IT?

1. **Dev environments setup**
2. **QA setup**
3. Code Architecture
4. Development
5. Product management

"One of the most powerful things that organizations can do is to enable development and testing to get environment they need when they need it"

- Gene Kim, author of "*The Phoenix Project*"

Data is the constraint

CIO Magazine Survey:

60% Projects Over Schedule

85% delayed waiting for data

Data is the Constraint

only getting worse

Gartner: Data Doomsday, 1/3rd IT in crisis

What Problems does the Data Constraint Cause ?

Operations

- Complex, manual data delivery takes days or even weeks
- Limited availability of test data to meet project team requirements
- High operational burden to ensure sensitive data is not exposed in non-production
- Limited resources: environments, storage

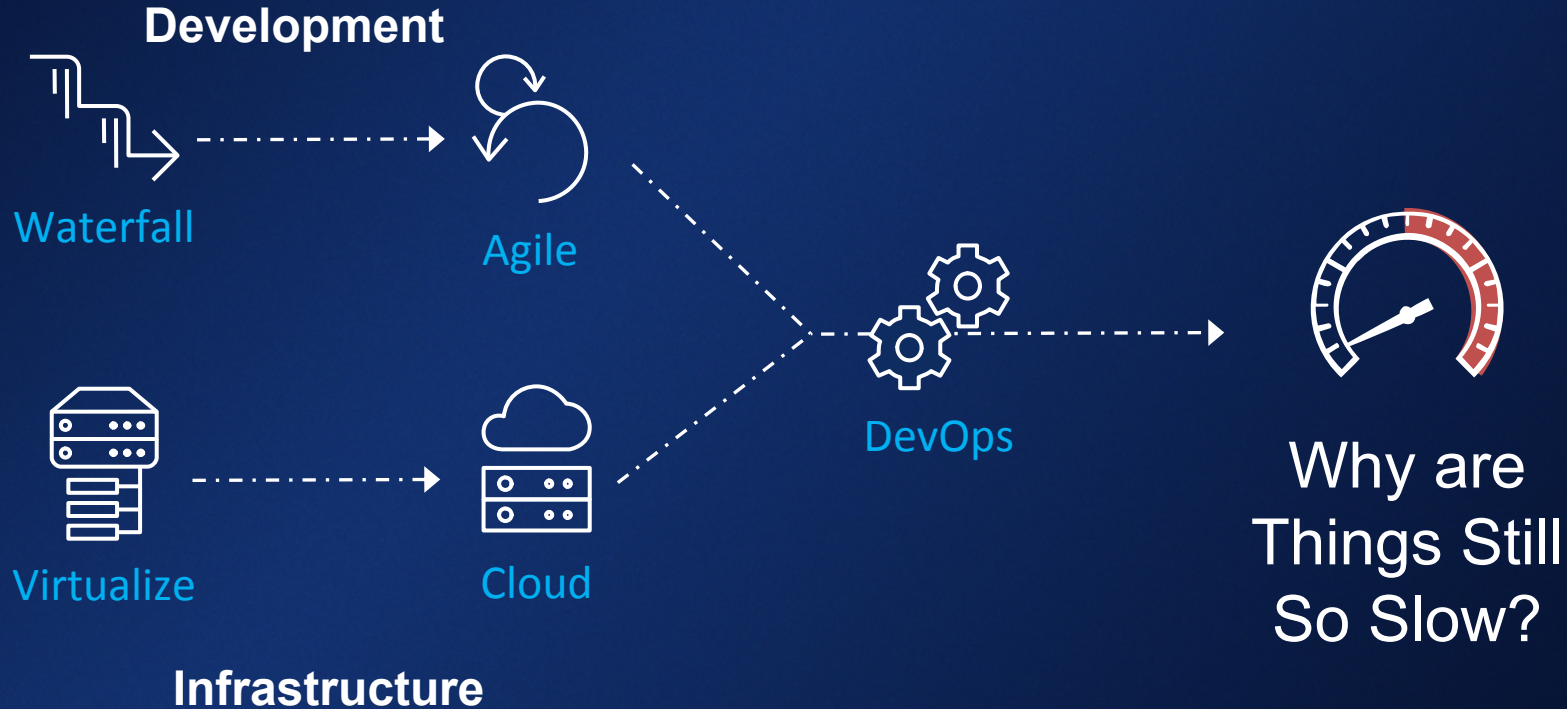## Increasing Complexity, Cost, and Risk

App Teams

- Unproductive wait-time during test cycles
- 20% of a tester's time spent resolving data-related issues
- Instances of sensitive data increases risk of data breach by up to 90%
- Complex release coordination and throttled throughput

More Data, More Owners, More Requests & Slower Time to Market

Development

Waterfall

Agile

DevOps

Infrastructure

Virtualize

Cloud

Why are Things Still So Slow?

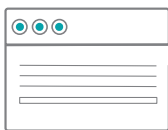# Data is Lagging Behind

Delphix Data Platform

Allows data to be delivered to every stakeholder at the speed, simplicity, and <u>security</u> required for business transformation.

The Delphix Dynamic Data Platform
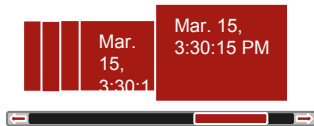
Collect

Databases

Applications &
File Systems

Storage = 10 TB

Control

Mask

Provision

Incremental
Sync

Backup, Retain

Mar. 15,
3:30:15 PM

Mar.
15,
3:30:1

Storage = 3 TB

Consume

Self-service virtual data on demand
Masked or Unmasked data

Storage = 0 TB

Bookmark   Refresh   Rewind   Branch   Integrate

Mask a Virtual Copy Once, Create Masked VDBs in Minutes

3 months ago   Last Monday   Today

10:27 A.M.   1:30 P.M.   5:07 P.M.

Next-gen data masking
- Easy to use
- Automatic profiling
- Referential integrity

Embedded native masking

DEV 1 to N

TEST 1 to N

UAT 1 to N

Cloud   Partners   On-Premises

Full, Virtual, Self-Service Capability

Provision   Refresh   Branch

Bookmark   Rewind   Synchronize

✓ Mask Once   ✓ Distribute Many   ✓ Refresh Anytime

Power Features and Data Governance Across Teams

Application Teams: **Power DevOps Data Features**

BRANCH
Create parallel development, test environments

REFRESH
Refresh to the latest data for high fidelity testing, analytics

INTEGRATE
Reset to run A/B tests, recover from errors and data loss

BOOKMARK
Mark release changes, business events

SHARE
Move environments to accelerate collaboration

IT Operations: **Full Security, Control, and Resource Allocation**

AUTOMATE
Provision apps, DBs with right data in minutes

MASK
Replace private, sensitive data with masked values

RETAIN
Store 30 days in space of one for restore, archive

REPLICATE
Replicate within or across sites, or   to the cloud

TRANSFORM
Convert Unix data into to Linux for modernization

Demonstration: Delivering Self-Service Data

JetStream and Timeflow – Your Best Friends…
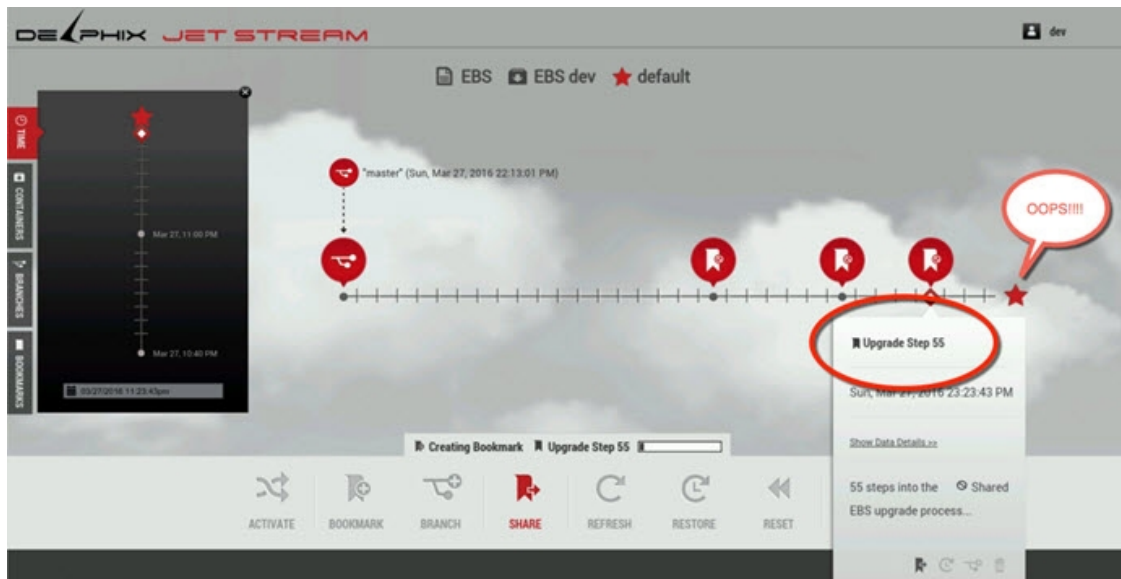
- Timeflow as seen in JetStream

- Data-as-a-Service on a red carpet!

- DBAs/Developers/QA Testers can rewind the entire system to a specific point-in-time or to a specific user-defined bookmark… 1 click and voila!

- Working on multiple branches and sharing datasets between groups

- ALL OF THAT done in seconds

On-Demand Access to Test Data
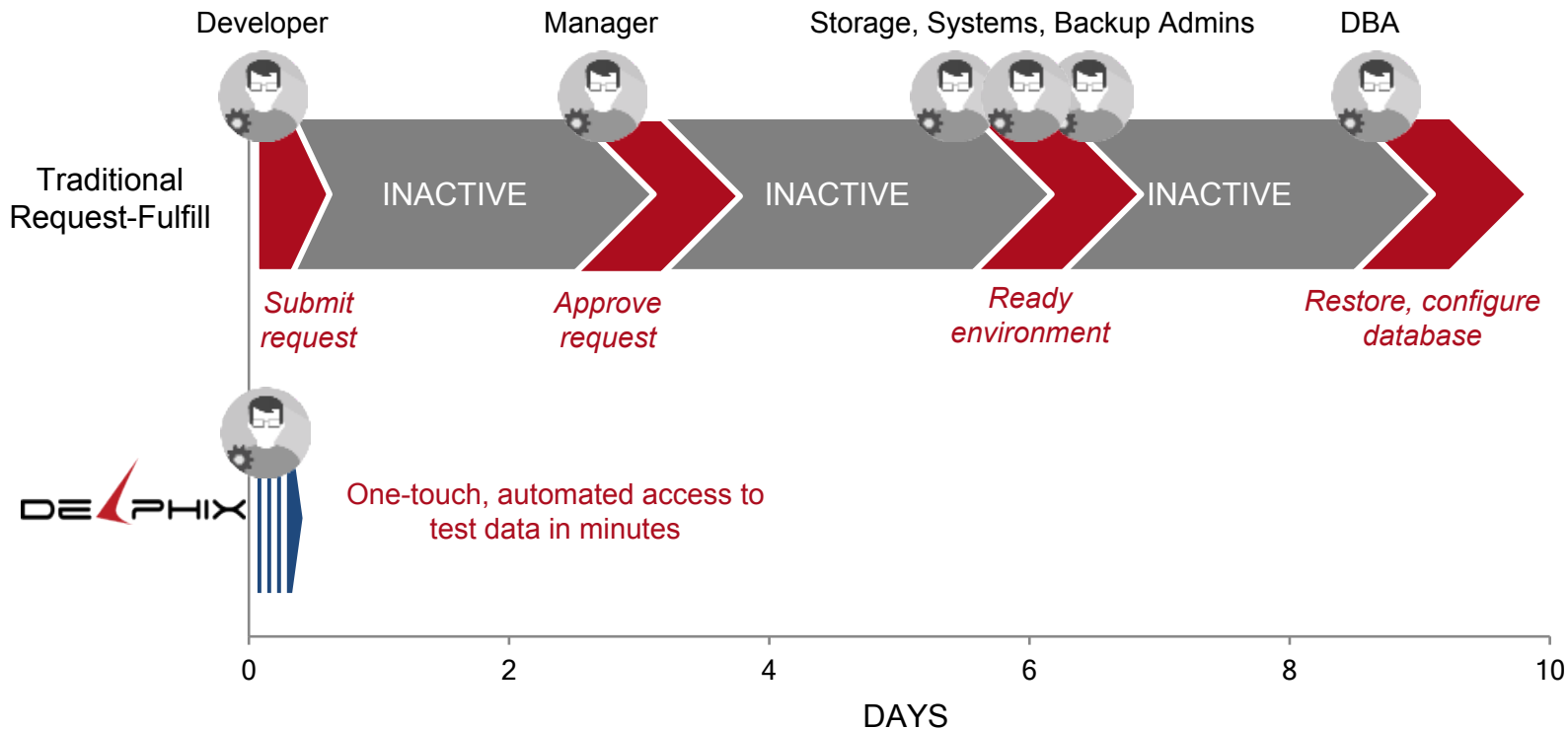
Improve Test Coverage with Bookmark, Reset

Physical Data

Virtual Data

Environment Reset

Environment Reset

Environment Reset

Test

Test

Test

Unproductive wait-time vs. testing using low-quality data

Bookmark, Reset

0   2   WEEKS   4   6
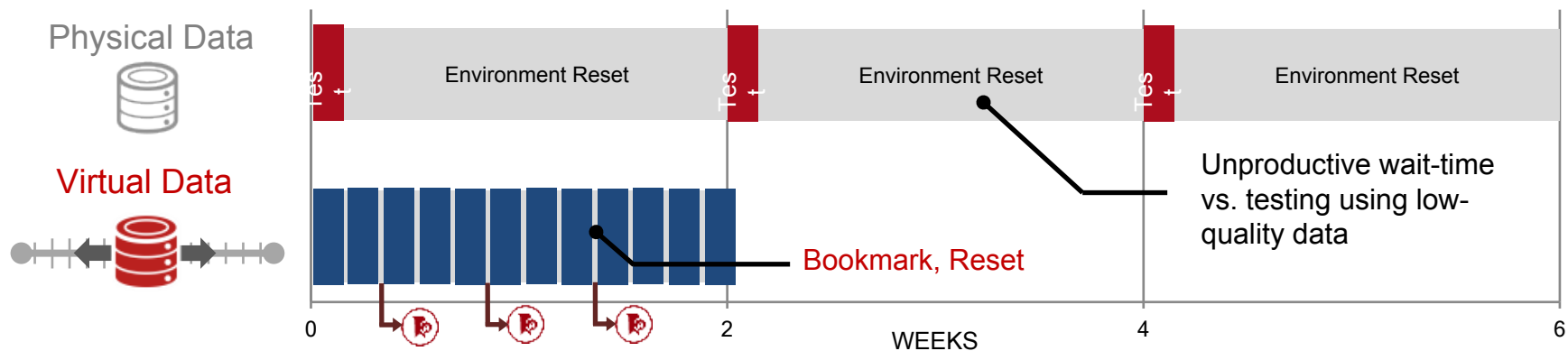
99% Less Downtime
Quickly bookmark, reset data to execute more testing in less time

Version Control
Orchestrate and align the right data with application code to improve test quality

Data Federation
Set up and rollback integrated test environments across multiple apps

Streamline end-to-end integration testing
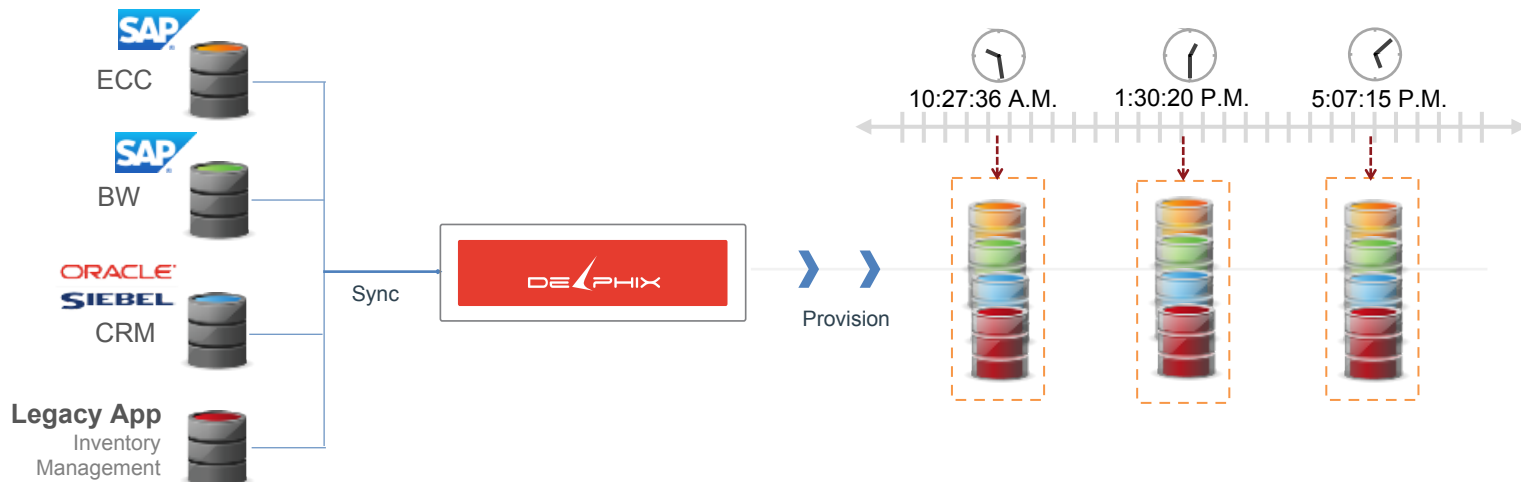
- Federate data to the same point in time across apps
- Enable faster releases, higher-quality testing

DevOps Readiness
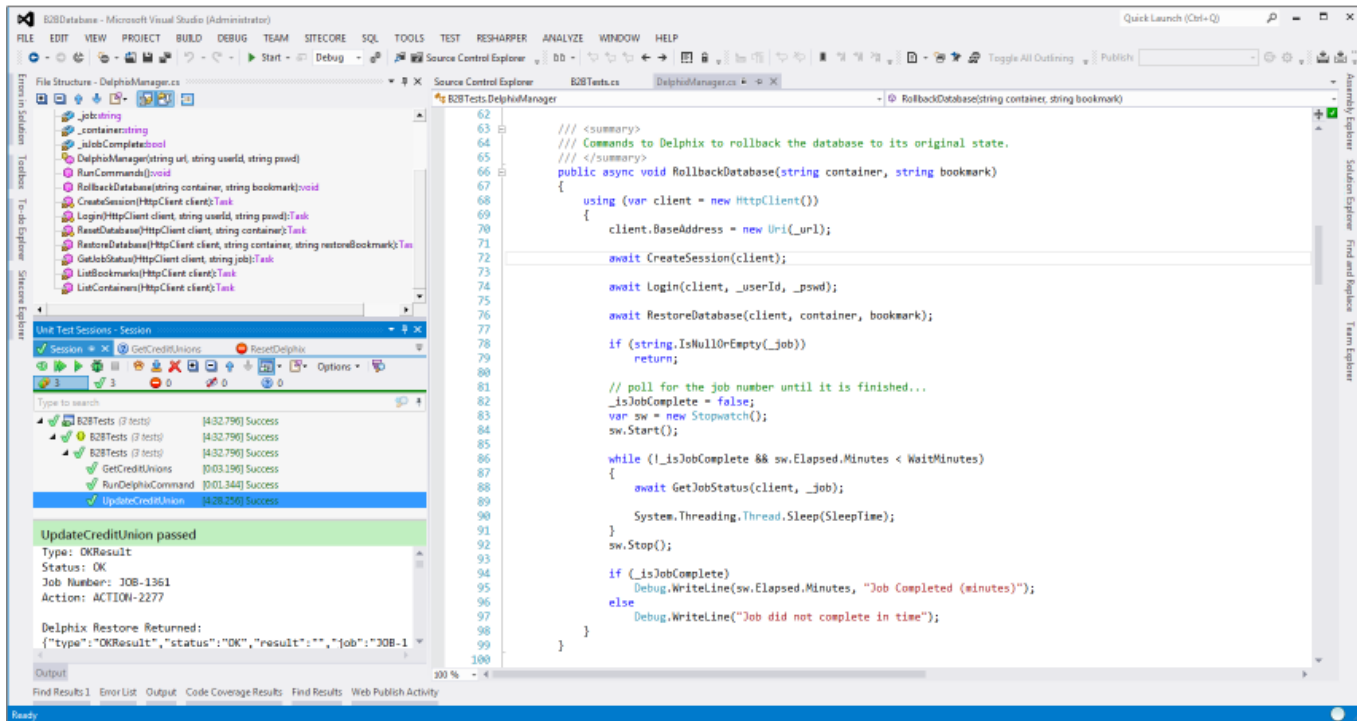
- Delphix APIs enable ready integration with existing DevOps tools and workflows
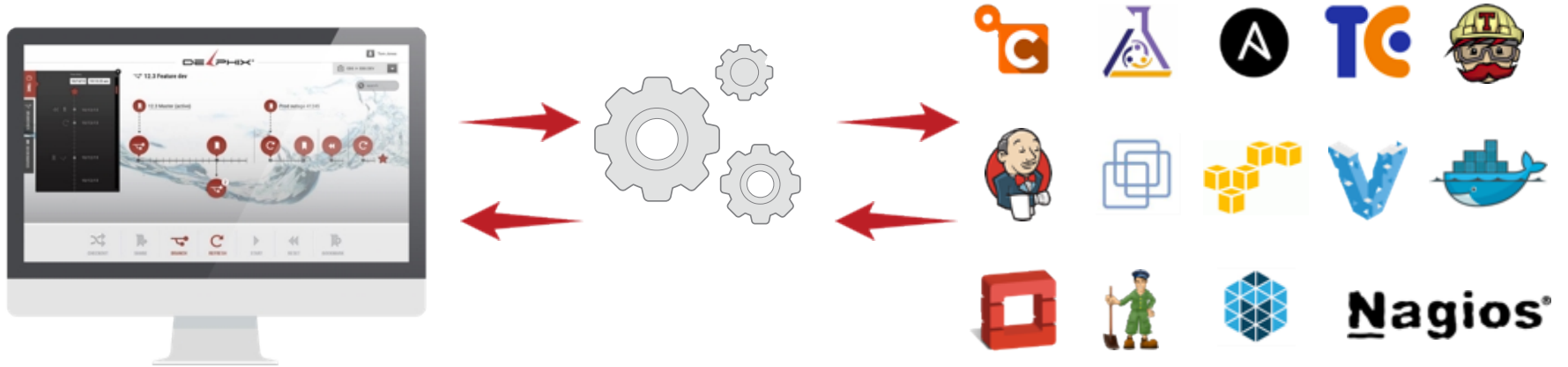
Demonstration: Delivering Self-Service Data
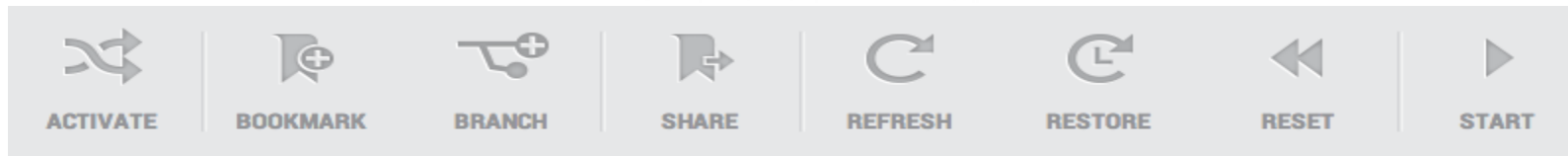
Regional Customer Sees Faster Time-to-Market

- Used Visual Studio to create automated test that depended on data being staged correctly

- Called Delphix from TFS Build to kick off database restore automatically

- Waited for restore to complete (4-7 minutes) before running automated test, guaranteeing same initial state every time.

Regional Customer Sees Upgrade Project Acceleration

Delphix Benefits for MIR

- ➢ Accelerated Productivity
  - ➢ Self-service for MarkitEDM data and code with bookmarks
  - ➢ Decoupled external and inner team dependencies
  - ➢ Empowered Project Team to operate in Agile and Continuous Delivery mindset

- ➢ Improved MarkitEDM Release Cycle Time
  - ➢ Provisioning Test environment
    - ➢ **Before** Delphix **40** hours every Sprint
    - ➢ **After** Delphix **3** hours every Sprint

2

Completed the MarkitEDM Application Upgrade and Migrations of 30+ databases from Old to New SQL Servers in _12 Weeks._

"Without Delphix and our new Agile Development methods, this type of effort would have been a project in itself, taking between 6-12 months."

"For an entire month before going live, we were pulling a fresh copy of production daily into our Continuous Integration environment using the Delphix API from TeamCity and performing the MarkitEDM application upgrade.   All of this was automated and really helped with our testing."

1. By removing dependencies we have reduced delivery of databases by **several** weeks
2. Business need for project teams to quickly provision test environments with data and code to enable agile development and testing
3. Reproduction of defect using Delphix shared bookmarks
4. Simplified execution and re-validation for Engineered Test Suite

Delphix: Data at the Speed of Business

QUALITY ★	SPEED ⏱	COST $
• Full, faithful, scalable data supporting test operations	• 1000x increase in test cycles through refresh + reset	• Delphix deploys on existing infrastructure
Result:	• 15-minute data standup	**Result:**
• Hit quality targets, avoid production bugs	**Result:**	• Revolutionize data delivery without further expense
	• Projects meet accelerated delivery deadlines	• Deploy without delay

StubHub!
» 30% revenue increase
» Monthly to daily releases
» 5 to 200 test cycles in existing systems

Walmart
» 10X test cycle acceleration
» Competitive edge in mobile, online

Resources

- Matthew.Moffat@delphix.com - Sales

- Todd.Blatti@delphix.com - Engineering

- Delphix.com

Your Data Has Never Done This Before

Data at the
Speed of Business

DELPHIX

Wells Fargo

Regulatory reporting – massive storage savings with significantly greater capabilities

- Expanded from 140 to 540 environments
- Increased online history from 1 to 90 days of history
- Reduced storage from 286 to 7 TB (98%)

Molina Healthcare

Deployed 14K masked VDBs, saving 5+ PB and reduced release cycles by 50%

"In my 4 years as CIO here at Molina, the best ROI of any technology investment that I've made has been with Delphix." – Rick Hopfer

Hewlett Pack Enterprise

Separated 450 SAP systems with <2 hours of downtime during go-live

"Delphix allowed us to complete the largest corporate separation in history, on time, with low risk." – Ralph Loura, CIO

HSBC

Deployed 300 VDBs in 3 days vs. 125 days.

"The demand internally is so large that we are trying our best to hold people off so we can keep up with the requests to use Delphix."

NIKE

Refresh data in 20 minutes vs. 12 hours.

"This quarter, we refreshed 1,638 times. On average, it takes 12 hours before and 20 minutes now to refresh. That's a savings of 40,956 hours this quarter."

JPMorganChase
2016 Hall of Innovation

"Delphix allows us to virtualize and mask data enabling our development teams to securely provision, refresh or rewind their data to any point-in-time. This is helping with our organization's transition to a more Agile and DevOps development methodology."

-Michael Urciuoli, CIO for J.P. Morgan Asset Management.

Delphix Value Domains

Development & Testing
- Reduce project/release cycles 50%
- Reduce data tasks from days/weeks to minutes
- Reduce non-prod storage by 90%
- Eliminate data risk in non-prod
- Debug with current production data
- Increase business velocity at reduced risk

Analytics, Modeling, & Reporting
- As much data as you want
- As many copies as you want
- Refreshed whenever you want – in minutes
- Eliminate batch job dependency and simplify ETL
- Fully masked or not – your choice

Regulatory Reporting
- Synchronized history across many sources – accessible to any point in minutes
- Significantly simplify data collection
- Massive savings in storage
- Significant data and history at your fingertips

Cloud, Big Data, & DC Migrations
- Migrate 50% faster with 20% volume
- Significantly reduced storage
- Eliminate data risk in non-prod
- Extract prod data anytime of day
- Seamless migration

Data Protection / Recovery
- Industry-changing simplicity
- Full-function, referential integrity
- Refresh any time and copy often
- Timeline of data allows for rapid recovery from data corruption and failed changes

If you're going to compete in the information economy, you need speed. You need Delphix.