What makes a *successful* Measurement Program?

- a case study

AEMES Conference Madrid

Presented by: Pam Morris

(CEO)

TOTAL METRICS

October, 2007

Presenter - Pam Morris

❑ CEO - TOTAL METRICS

❑ Consulting, Training Tools and Standards
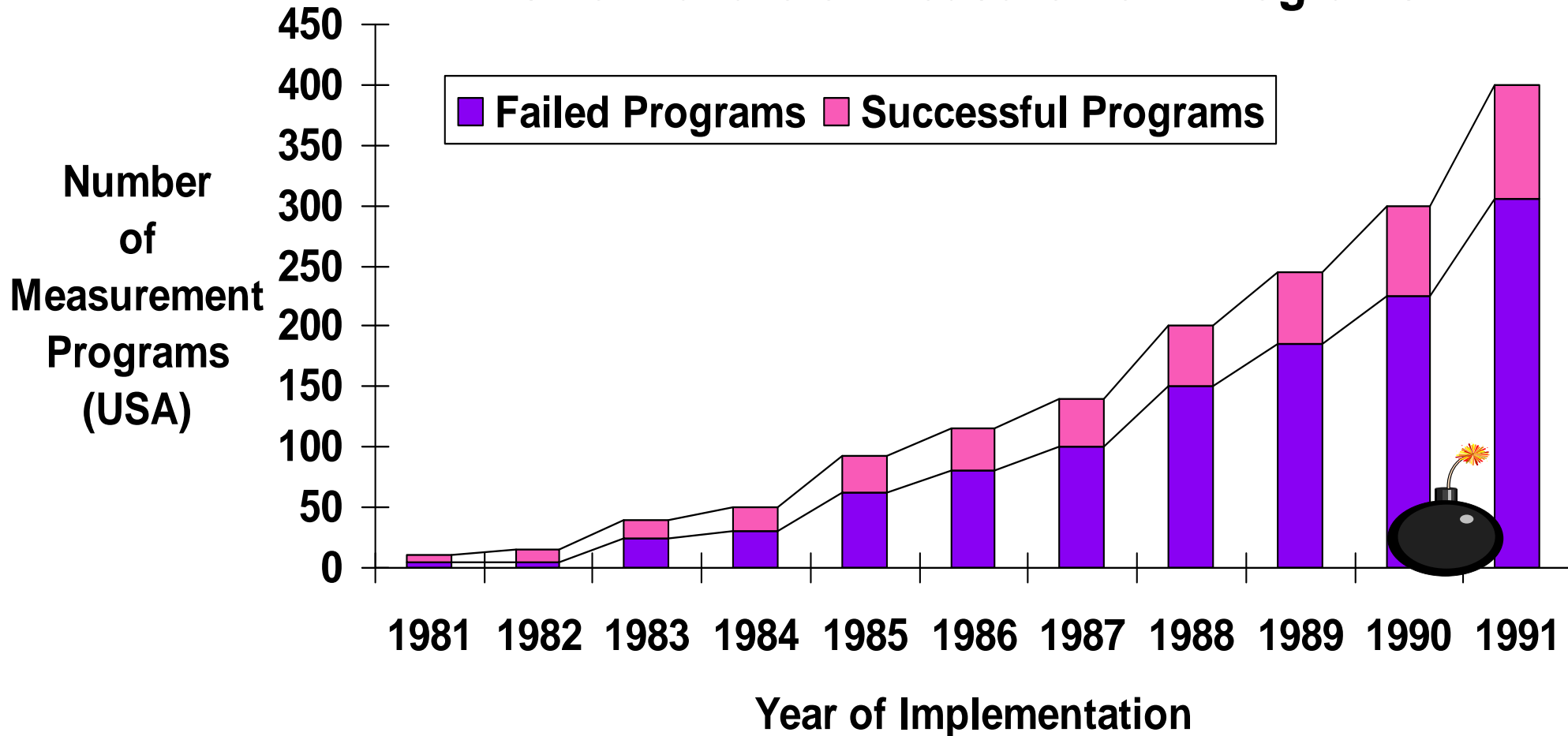 o Certified IFPUG (CFPS, CSMS - 3), COSMIC-FFP

❑ Committee Member of:
  o Australian Software Metrics Association (ASMA) – Executive (1991 -)
  o ISBSG – Executive (2000 -)
  o International Function Point User Group (IFPUG) (1993 - 2000)
  o COSMIC-FFP Core Committee – (1997 -)
  o International Standards Organisation (ISO) - WG12 (1993 – 2007)
  o Standards Australia - IT15 (1993 -)

Has anything changed?

◆ *"80% of all measurement programs fail "*

Source : Howard Ruben Associates 1994

Risk of Failure of Measurement Programs — Number of Measurement Programs (USA) vs Year of Implementation (1981–1991), stacked bar chart of Failed Programs and Successful Programs.

Overview of Topics

◆ **Background**

◆ **Measurement Process**

◆ **Lessons Learned**

◆ **Critical Success Factors**

Background

◆ Australian Government Department

◆ Large Legacy Application - ~14,000 fps

◆ Mid-range – Cool:Gen, Java

◆ 60 developers

◆ Initial Objective : Verify improvements gained by Re-factoring activity
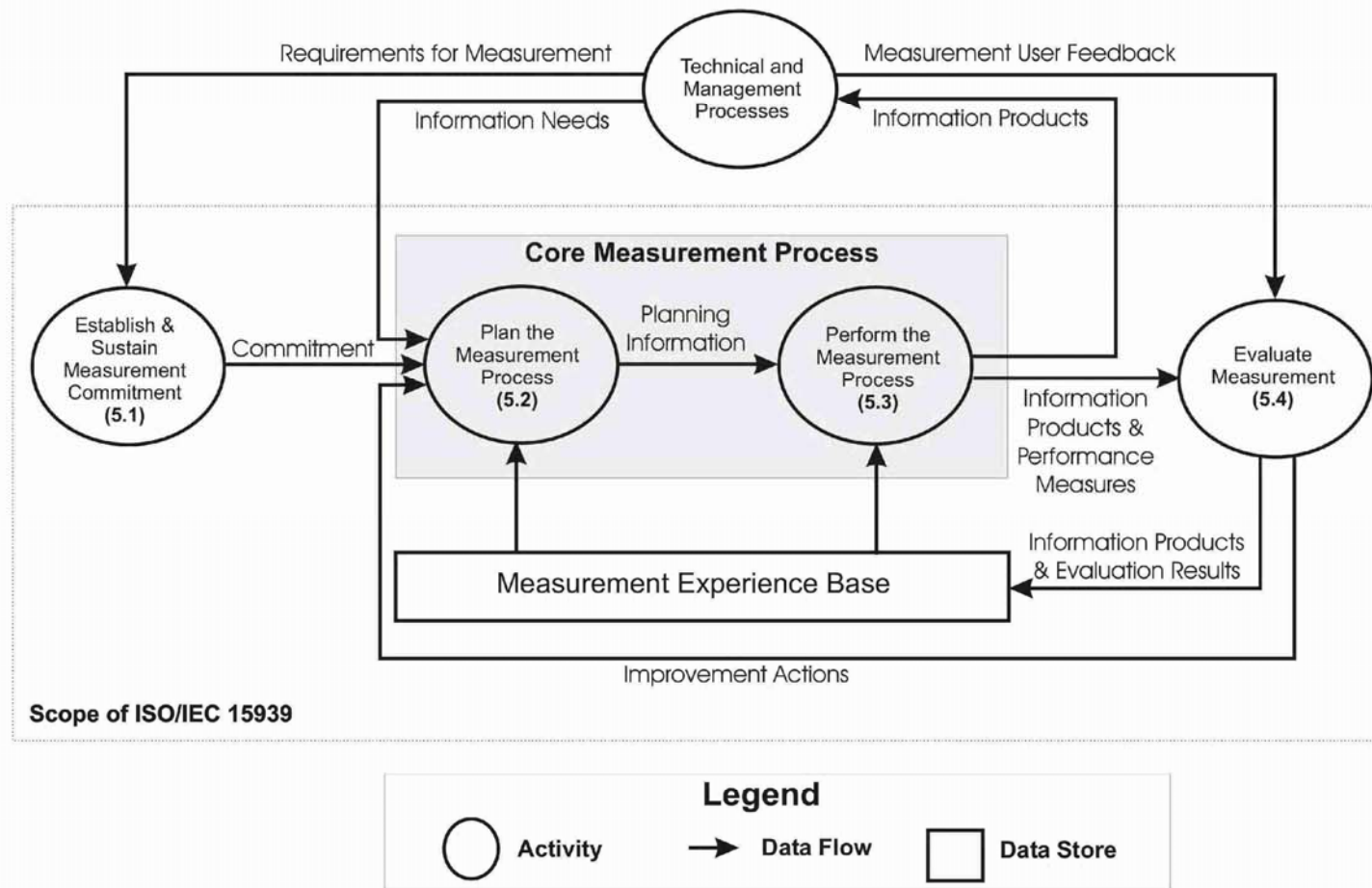
Measurement Process – ISO/IEC 15939:2007

Requirements for Measurement

Information Needs

Measurement User Feedback

Technical and Management Processes

Information Products

Core Measurement Process

Establish & Sustain Measurement Commitment (5.1)

Commitment

Plan the Measurement Process (5.2)

Planning Information

Perform the Measurement Process (5.3)

Evaluate Measurement (5.4)

Information Products & Performance Measures

Measurement Experience Base

Information Products & Evaluation Results

Improvement Actions

Scope of ISO/IEC 15939

Legend

○ Activity   → Data Flow   □ Data Store

Figure 1: Measurement Process Model.

1. Establish and Sustain Measurement and Management Commitment

◆ Management had clear stated objectives

◆ 4 Year commitment

◆ Buy in from CIO to Project Team Leaders

2. Plan the Measurement Process

◆ Workshops to agree:

➤ KRA, KPIs

➤ Report Templates

➤ Data Collection Templates

➤ Tools

Metrics Consultant Resource:

◆ 1 consultant

◆ Effort = 22 days

◆ Duration = 1 Calendar Month

Client Resource:

◆ 4 Management

◆ Effort = 1 ½ day workshops + Review 2 Drafts

◆ Duration = 1 Calendar Month

Reporting Structures

No.	Name	Key Result Area	Report Level	IT Steering Committee	TeamLeaders Project Board	QC Mngment
				Target Audience		
Main Reports						
1	ARLS Productivity and Quality	Cost & Quality	Release / Cumulative	√	√	
2	ARLS Productivity and Release Size	Cost	Release / Cumulative	√	√	
3	ARLS Release Quality and Testing Effectiveness	Quality	Release / Cumulative	√	√	√
4	ARLS Baseline Growth	Cost (Investment)	Application / Cumulative	√	√	
Supplementary Reports						
5	ARLS Project Productivity and Quality	Cost & Quality	Project / 6 month snapshot		√	√
6	ARLS Project Quality and Testing Effectiveness	Quality	Project / 6 month snapshot		√	√
7	ARLS Analysis of Defects – by Severity	Quality	Release / 6 month snapshot		√	√
8	ARLS Analysis of Defects – by Source of Origin	Quality	Release / 6 month snapshot		√	√
9	ARLS Development Stage Analysis	Quality / Cost	Project / 6 month snapshot		√	√
10	ARLS Time Spent in Testing	Cost / Quality	Release / 6 month snapshot		√	√
11	ARLS Rework Analysis - Summary	Cost / Quality	Release / Cumulative		√	√
12	ARLS Rework Analysis - Detail	Cost / Quality	Project / 6 month snapshot		√	√
13	ARLS Maintenance Intensity	Cost	Application Cumulative		√	√

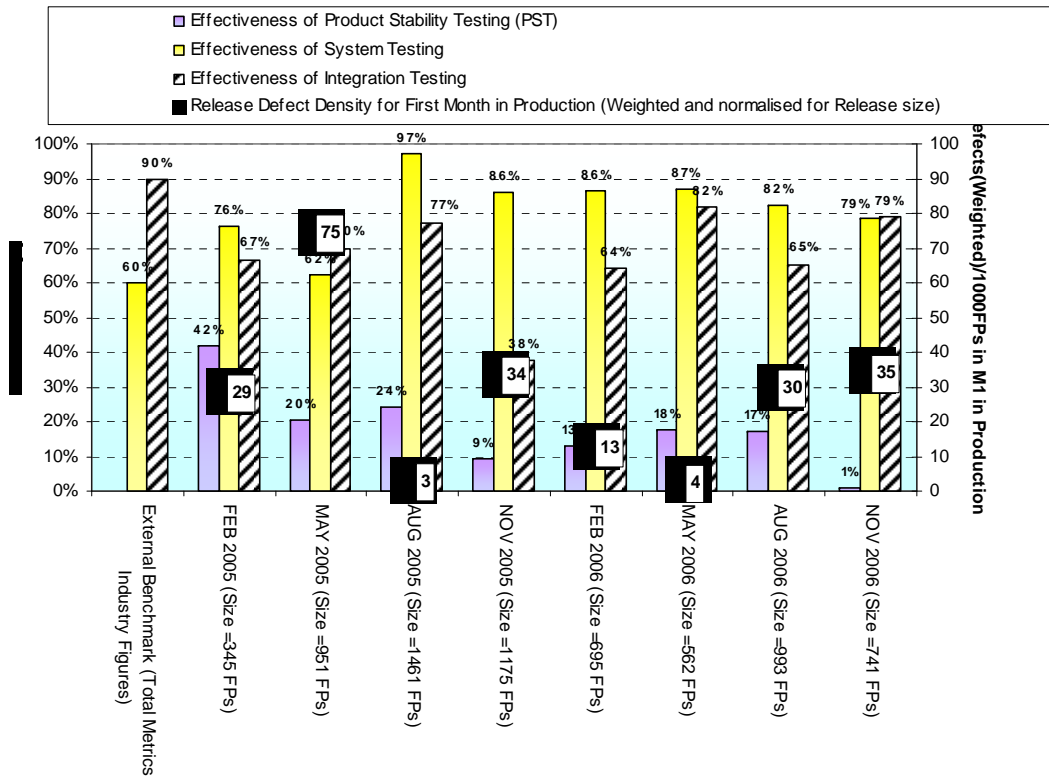# Report Templates

◆ Each Report had agreed:

> ➤ Purpose

> ➤ Target Audience

> ➤ Frequency / Level

> ➤ Rules for Calculation

> ➤ Description :

>> ✴ how to read the report

>> ✴ What it was demonstrating

>> ✴ the types of decisions it would support

Report Templates

ARLS Release Test Effectiveness

Note: Testing Effectiveness compares the number of defects found at a particular stage of testing against how many defects were actually le

- ☐ Effectiveness of Product Stability Testing (PST)
- ☐ Effectiveness of System Testing
- ◪ Effectiveness of Integration Testing
- ■ Release Defect Density for First Month in Production (Weighted and normalised for Release size)

◆ *Eg Description*

◆ This report shows the overall quality of the *ARLS* development process since the degree to which defects are released into production are a good indication of the maturity of software development.

Data Collection Templates

◆ 5 Base Measures and Tools Agreed:

➢ Functional Size (fps)
  * IFPUG 4.2
  * *SCOPE Project Sizing Software*™

➢ Effort (hours)
  * ISBSG Definitions Level 2
  * *NIKU*™

➢ Defects (number)
  * origin, severity
  * ISBSG Customised
  * *Test Track Pro* ™

➢ Duration (Calendar Days)
  * ISBSG Definitions
  * *NIKU*™ Rules for Calculation

➢ Full-time Equivalents (people)
  * ISBSG Definitions

3. Perform Measurement Process

◆ Establish Baseline

Metrics Consultant Resource:

- ◆ 1 consultant
- ◆ Effort = 33 days
- ◆ Duration = 2 Calendar Months

Client Resource:

- ◆ 8 application experts
- ◆ Effort = ~1/2 day each
- ◆ Duration = 2 Calendar Months

◆ Ongoing Measurement

➢ ~ 6 projects every 3 month Release (846fps)

Metrics Consultant Resource:

- ◆ 1 consultant
- ◆ Effort = 5 days
- ◆ Duration = 1 Calendar week

Client Resource:

- ◆ Project Teams
- ◆ Effort = ?
- ◆ Duration = 3 Calendar Months

3. Perform Measurement Process

◆ Analysis of the Results – 52 KPIs

Metrics Consultant Resource:

- ◆ 1 consultant
- ◆ Effort = 5 days
- ◆ Duration = 1 Calendar week

Client Resource:

- ◆ 1 Metrics Analyst
- ◆ Effort = 10 days
- ◆ Duration = 1 Calendar Month

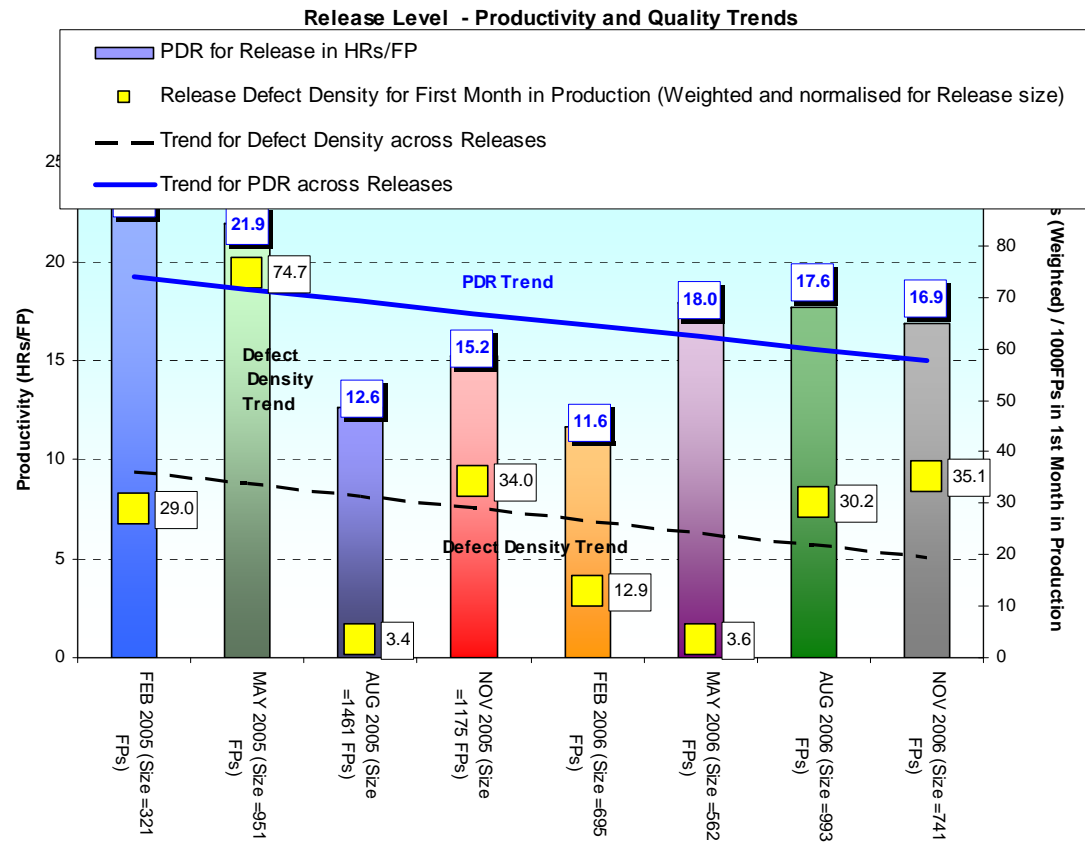◆ Reporting the Results

➢ Benchmark Report (6 monthly) – 100 pages

Metrics Consultant Resource:

- ◆ 1 consultant
- ◆ Effort = 10 - 15 days
- ◆ Duration = 3 Calendar weeks

Client Resource:

- ◆ Management Reviews
- ◆ Effort = 1 days
- ◆ Duration = 1 Calendar day

4. *Feedback into Technical and Management Processes*

Release Level - Productivity and Quality Trends

Legend:
- PDR for Release in HRs/FP
- Release Defect Density for First Month in Production (Weighted and normalised for Release size)
- Trend for Defect Density across Releases
- Trend for PDR across Releases

Productivity (HRs/FP) — vertical axis (left)

Defects (Weighted) / 1000FPs in 1st Month in Production — vertical axis (right)

Release	PDR	Defect Density
FEB 2005 (Size =321 FPs)	21.9	29.0
MAY 2005 (Size =951 FPs)	21.9	74.7
AUG 2005 (Size =1461 FPs)	12.6	3.4
NOV 2005 (Size =1175 FPs)	15.2	34.0
FEB 2006 (Size =695 FPs)	11.6	12.9
MAY 2006 (Size =562 FPs)	18.0	3.6
AUG 2006 (Size =993 FPs)	17.6	30.2
NOV 2006 (Size =741 FPs)	16.9	35.1

PDR Trend

Defect Density Trend

4. Feedback into Technical and Management Processes

◆ Product Quality

➢ Observations
- ✴ Most defects originated in Build phase
- ✴ Testing was introducing defects
- ✴ Testing efficiency was below industry standard
- ✴ Time spent early life cycle was below industry standard
- ✴ Large variability between projects

◆ Product Quality

➢ Improvements Introduced

* Peer Reviews
* Formal Unit Test process
* Focus on System Testing
* Formal Requirements Management and Design Process

4. Feedback into Technical and Management Processes

◆ Productivity

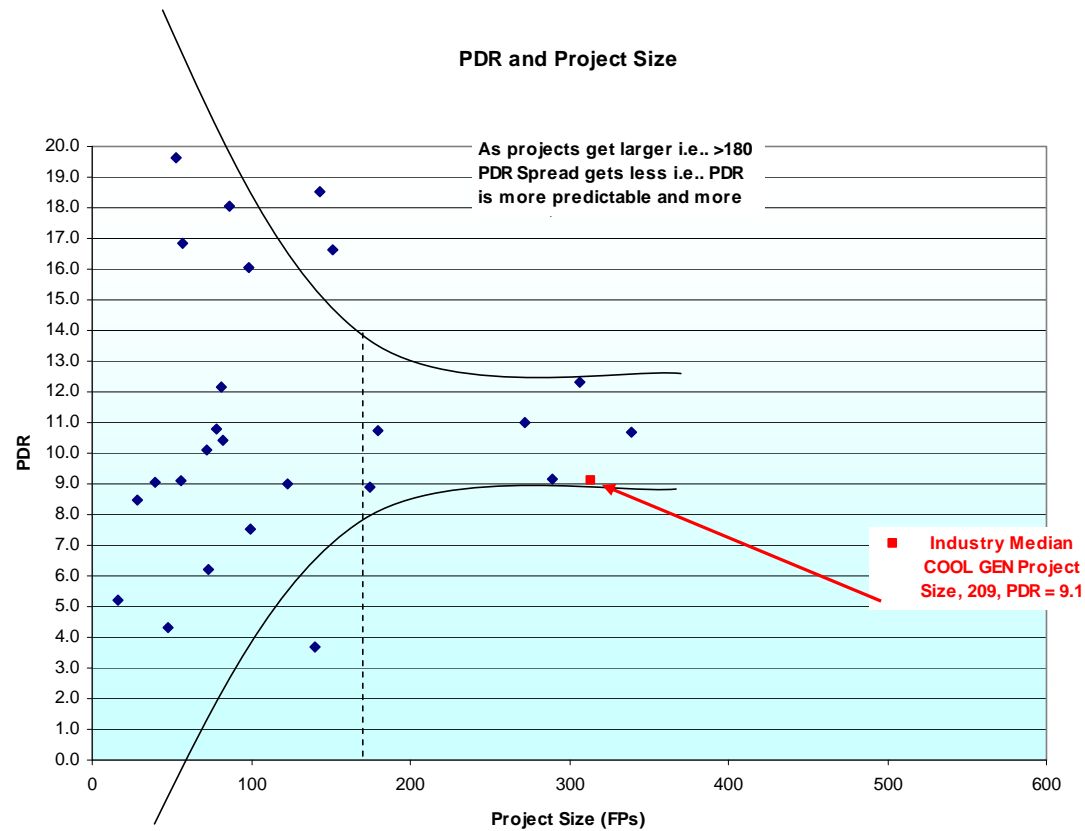➤ Observations

✴ Less productive than Industry

✴ Small projects (<100fps):

– have lower productivity
– Small projects behave unpredictably.

✴ Larger Projects (>250 fps) took longer users optimum 12 months

✴ FP size gave accurate early life cycle estimates
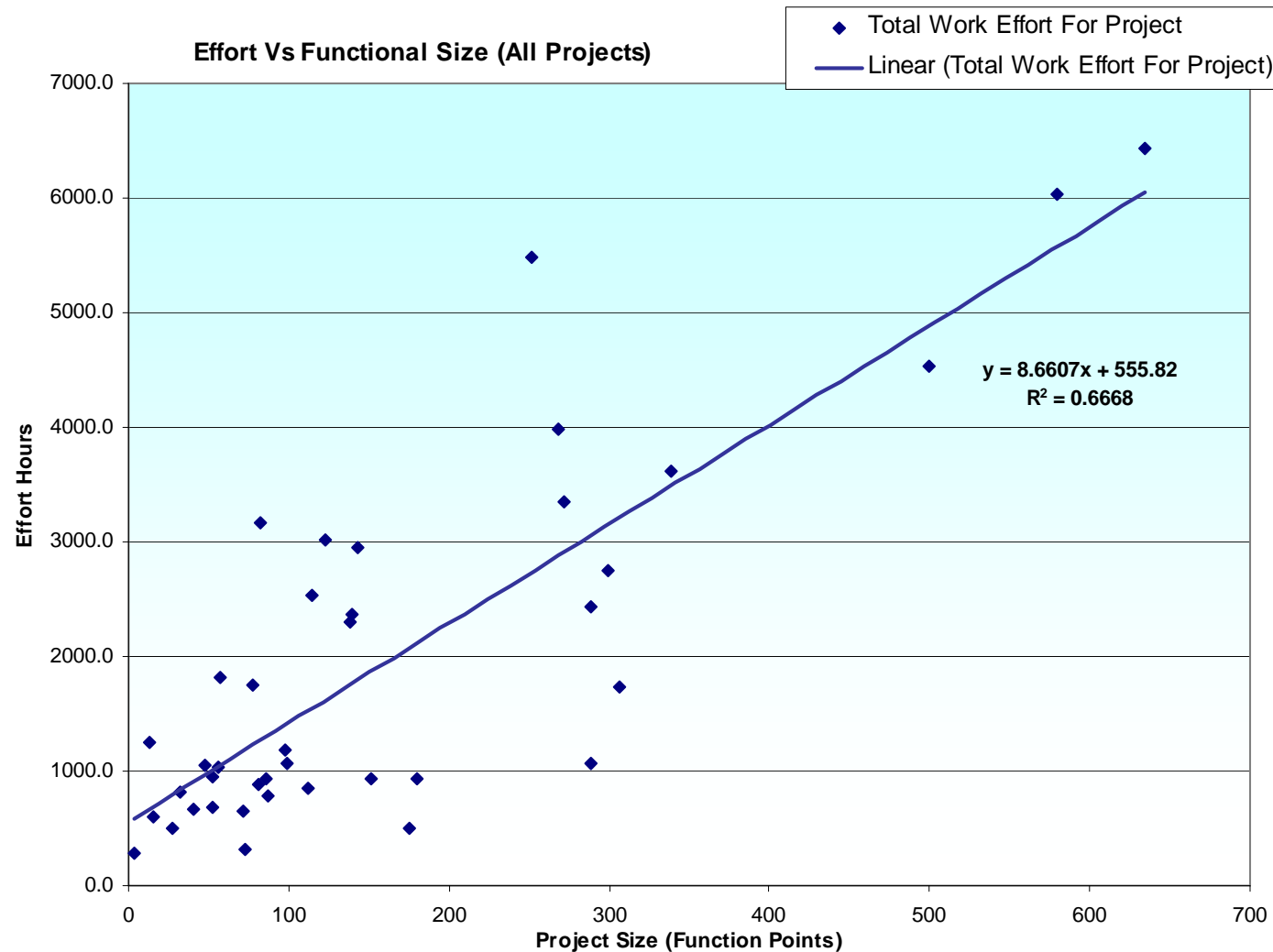
✴ Large variability between projects

Less Productive than Industry Median

Position	Project Median PDR Comparison to Industry by Release								Industry Values (R10 - 2007)		
	Feb-05	May-05	Aug-05	Nov-05	Feb-06	May-06	Aug-06	Nov-06	Cool:GEN	4 GL Projects	Case Tools
Minimum value									2.7	0.9	1.8
Top 25% of productivity					7.5				6.8	3.7	6.5
Median rate		12.2	12.1				10.1		9.1	6.7	14.4
Bottom 25% of Productivity	21.1			16.8		18.1		23.4	12.5	12.4	30.0
Maximum Value									56.1	40.5	80.7
Number in sample	2	4	8	6	7	7	5	4	28	89	81

Small Project are more unpredictable

PDR and Project Size

As projects get larger i.e.. >180 PDR Spread gets less i.e.. PDR is more predictable and more

Industry Median COOL GEN Project Size, 209, PDR = 9.1

PDR

Project Size (FPs)

FP Size has good correlation with effort

Effort Vs Functional Size (All Projects)

Legend:
- ◆ Total Work Effort For Project
- — Linear (Total Work Effort For Project)

$y = 8.6607x + 555.82$
$R^2 = 0.6668$

Y-axis: Effort Hours
X-axis: Project Size (Function Points)

Estimated FP Size Produced Accurate Effort Estimates

ARLS Release Effort Estimates Vs Actual Release (Project) Effort in Hours

Release Effort Hours (Development)

NOV 2005
14,702

FEB 2006
5,969

MAY 2006
8,426

AUG 2006
11,862

NOV 2006
10,366

■ Actual Effort (excludes Release Management) (+10% error bars)

— Predicted Effort hours using DRIVES (Release Level) median PDR

2nd Benchmark   3rd Benchmark   4th Benchmark

© Total Metrics 2007
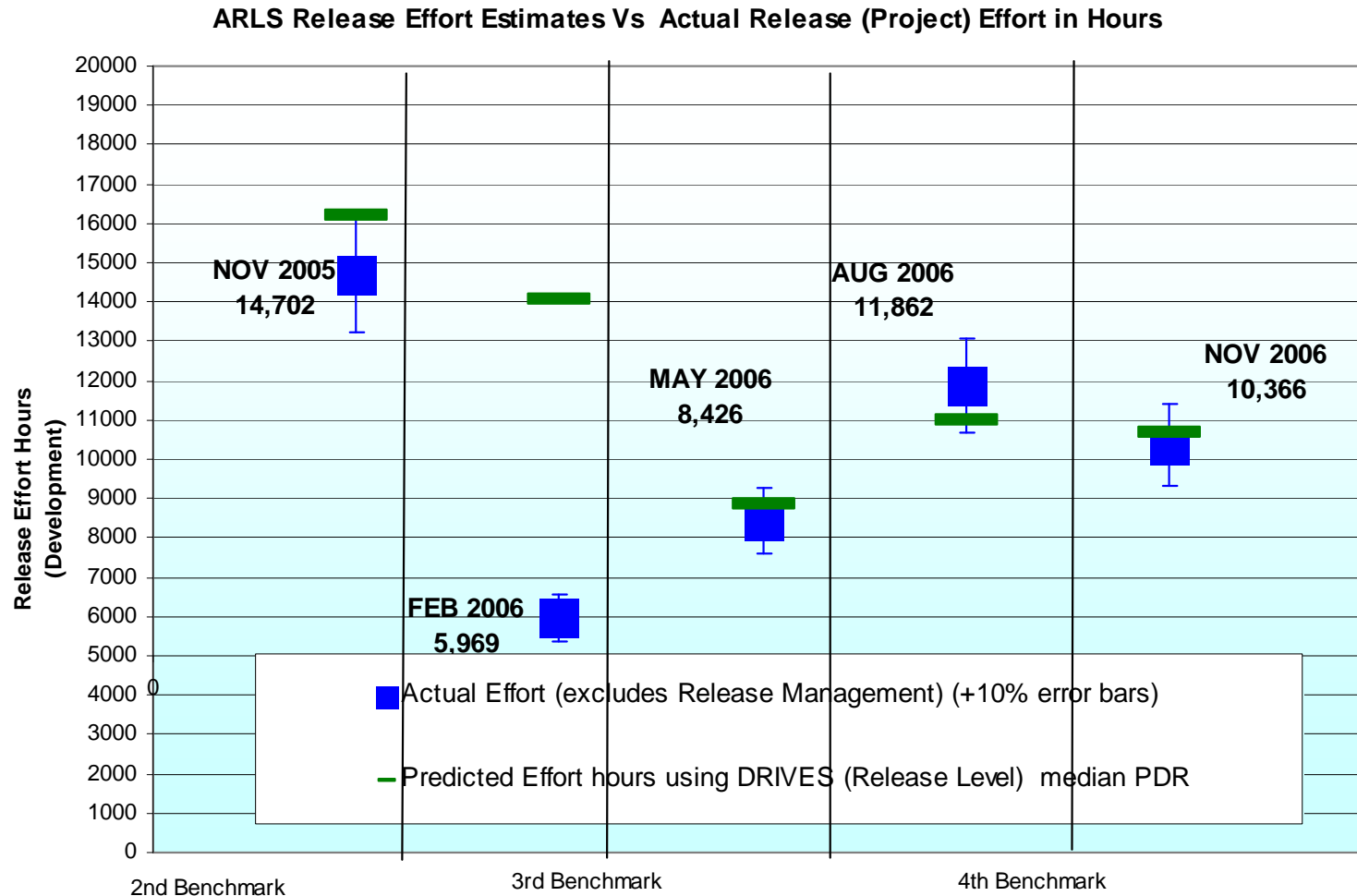
Only 33% of Projects delivering New functionality to the Business and Net Growth is decreasing

Release Functional Impact Analysis - Summary

Nett Growth For Release = (Added Fps)+ (Changed After size - Changed Before Size) - (Deleted)
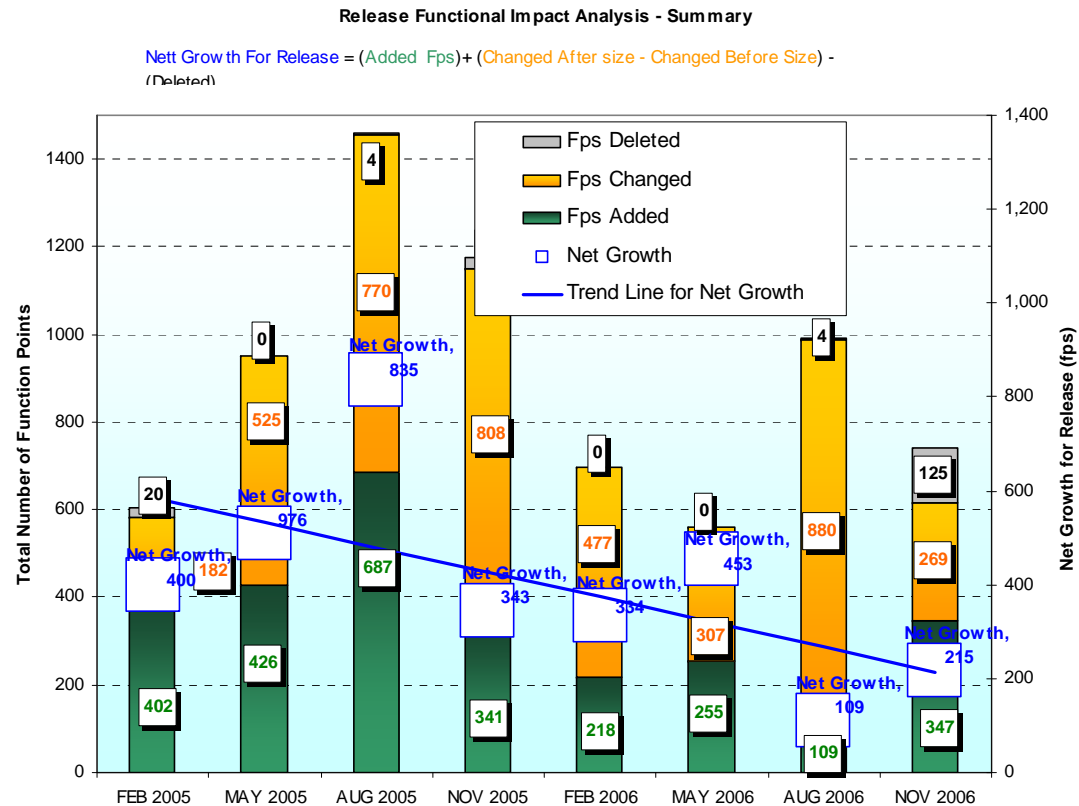
5. Evaluate Measurement

◆ Metrics Review Workshop – 2 hours

Metrics Consultant Resource:

- ◆ 1 consultant
- ◆ Effort = 2 hours
- ◆ Duration = 1 Calendar day

Client Resource:

- ◆ 5 Management team
- ◆ Effort = 1 day
- ◆ Duration = 1 Calendar day

◆ Implementing Changes

➢ Data Collection and Recording

Metrics Consultant Resource:

- ◆ 1 consultant
- ◆ Effort = 3 days
- ◆ Duration = 1 Calendar week

Client Resource:

- ◆ Metrics Analyst + Training
- ◆ Effort = 5 days
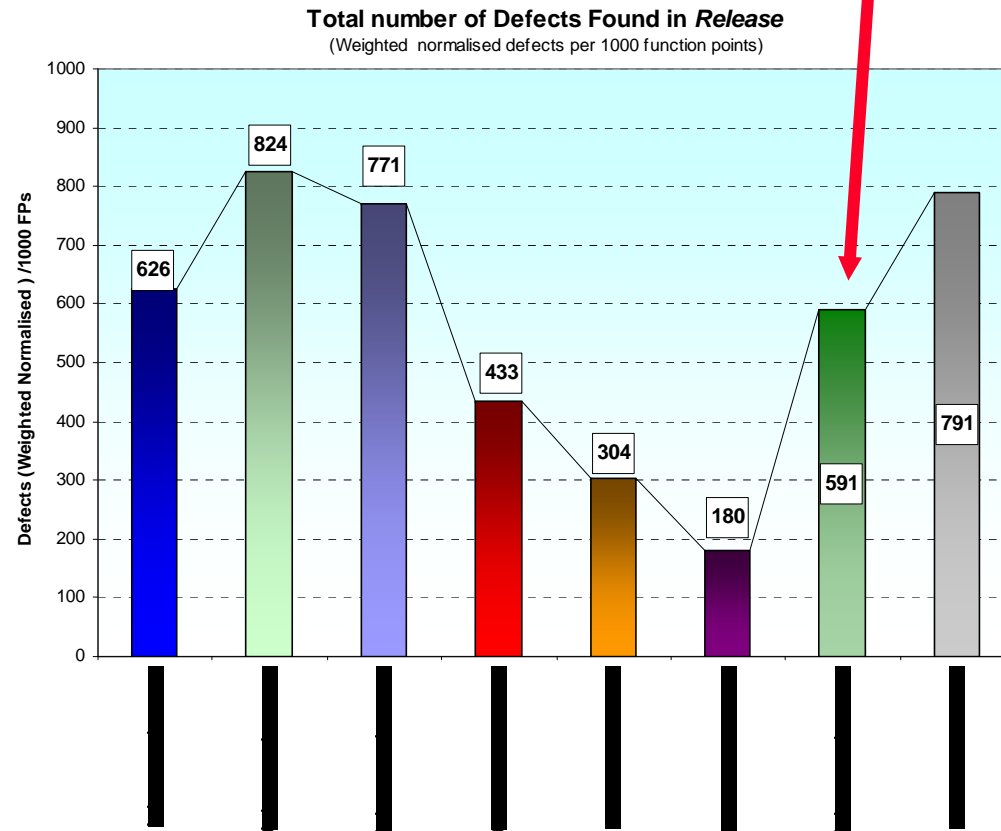- ◆ Duration = 1 Calendar month

Changes Introduced

◆ Defects

➢ All defects now captured – early life cycle

➢ Unit Testing defects now captured accurately

➢ Defects now allocated correctly to phase
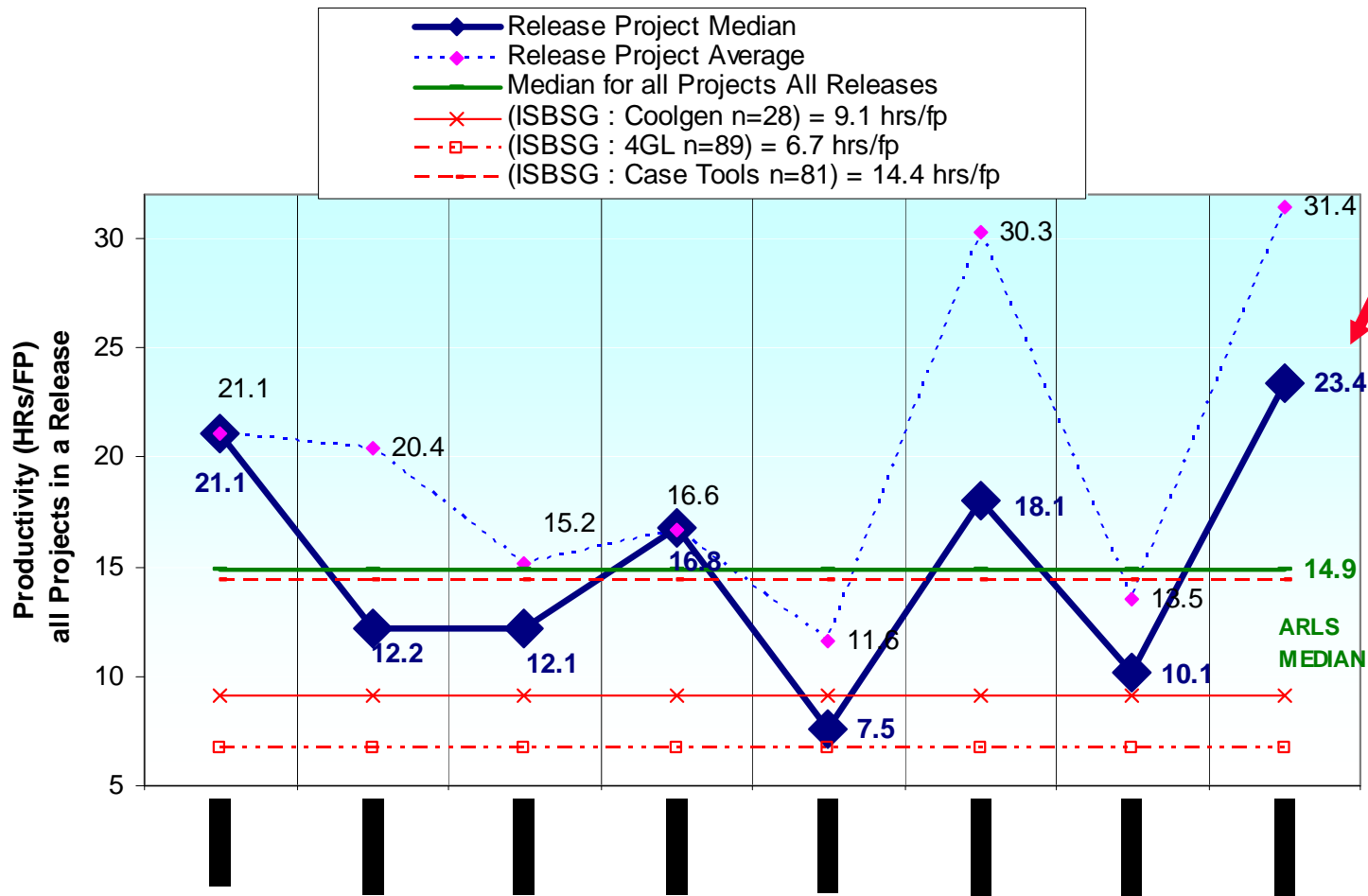
➡ 4th Benchmark more defects being reported

◆ Effort

➢ QC Effort now allocated to the project not the Release overhead.

➡ 4th Benchmark higher Project PDR being reported

More Defects being Reported

Total number of Defects Found in *Release*
(Weighted normalised defects per 1000 function points)

Lower Project Productivity (higher PDR) Reported

Comparison ALL *PROJECT* PDR to Industry Medians

Legend:
- Release Project Median
- Release Project Average
- Median for all Projects All Releases
- (ISBSG : Coolgen n=28) = 9.1 hrs/fp
- (ISBSG : 4GL n=89) = 6.7 hrs/fp
- (ISBSG : Case Tools n=81) = 14.4 hrs/fp

Y-axis: Productivity (HRs/FP) all Projects in a Release

Data points (Median): 21.1, 12.2, 12.1, 16.6, 7.5, 18.1, 10.1, 23.4
Data points (Average): 21.1, 20.4, 15.2, 16.8, 11.5, 30.3, 13.5, 31.4

ARLS MEDIAN: 14.9

X-axis: Benchmarking Release

Critical Success Factors

◆ Formal Process
- ➤ Clear Stated objectives
- ➤ Vision – long term commitment
- ➤ Adequate Budget and Resources
- ➤ Used skilled Metrics personnel
- ➤ Used specialist tools for FPA and outsourced counting

◆ Management
- ➤ Realistic expectations
- ➤ All levels interested, results are shared
- ➤ Acts on the results
- ➤ Open to change
- ➤ Sees bad news as an opportunity
- ➤ Measurement is viewed as important

News Flash - May 2007

True Measure of Success!

◆ Other IT Areas want what they have got!

◆ 8 other Applications want to be involved and get what the ARLS team are getting!
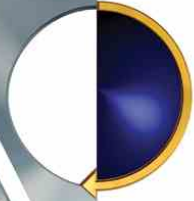
At Last Success !

Total Metrics Pty Ltd
667 Burke Road
Camberwell
Victoria   3124   Australia

Phone   +613 9882 7611

Fax   +613 9882 7633

admin@Totalmetrics.com
www.totalmetrics.com