

Continuous Delivery for Linux/Windows/Hadoop

Wisely Chen

Agenda

- Background
- Problem
- Solution
- Demo
- Q & A

Who I am



- Wisely Chen (thegiive@gmail.com)
- Release manager of Yahoo! Taiwan shopping and data team
- Enjoys promoting open source tech
 - Hadoop Summit 2013 San Jose
 - Ruby and Rails (All in Taiwan): Coscup 2006, Ubisunrise 2007, OSDC 2007
 - Release Practice (All in Taiwan): Webconf 2013, Coscup 2012,2013, PHPConf 2012 , RubyConf 2012

Team



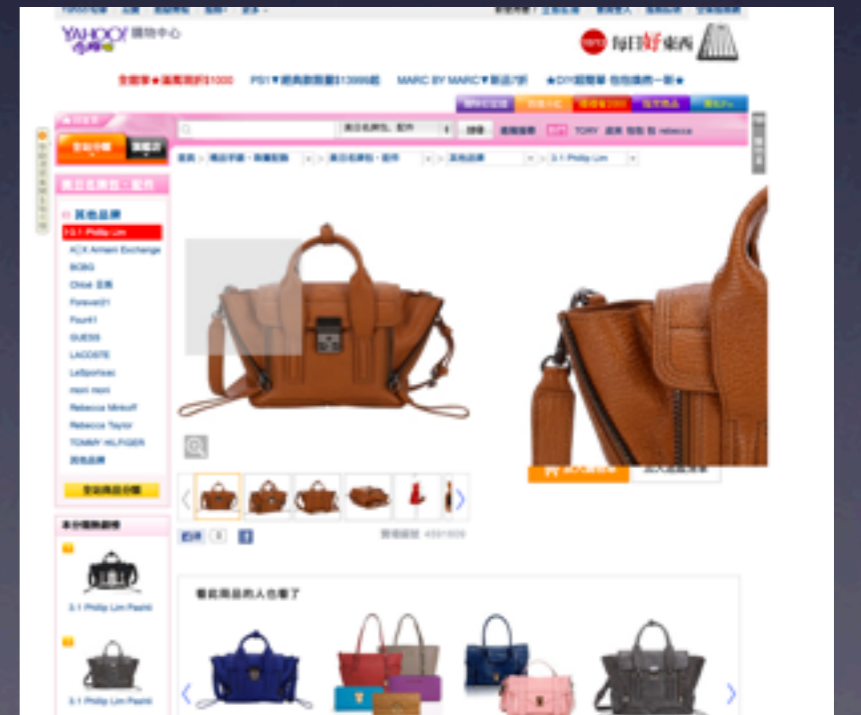
Sanching Yu
Windows



Neal Lee
Hadoop

Yahoo! Taiwan Shopping

- Top B2C site in Taiwan
- Monday.com acquired in 2008
- 8 engineering teams
 - 15K files changed monthly
 - 100K lines of code changed monthly



Platform

Frontend (Linux)

Catalog

Payment

Member

Mobile

Backend (Window)

Cash Flow

Supplier

Logistic

CMS

Data (Hadoop)

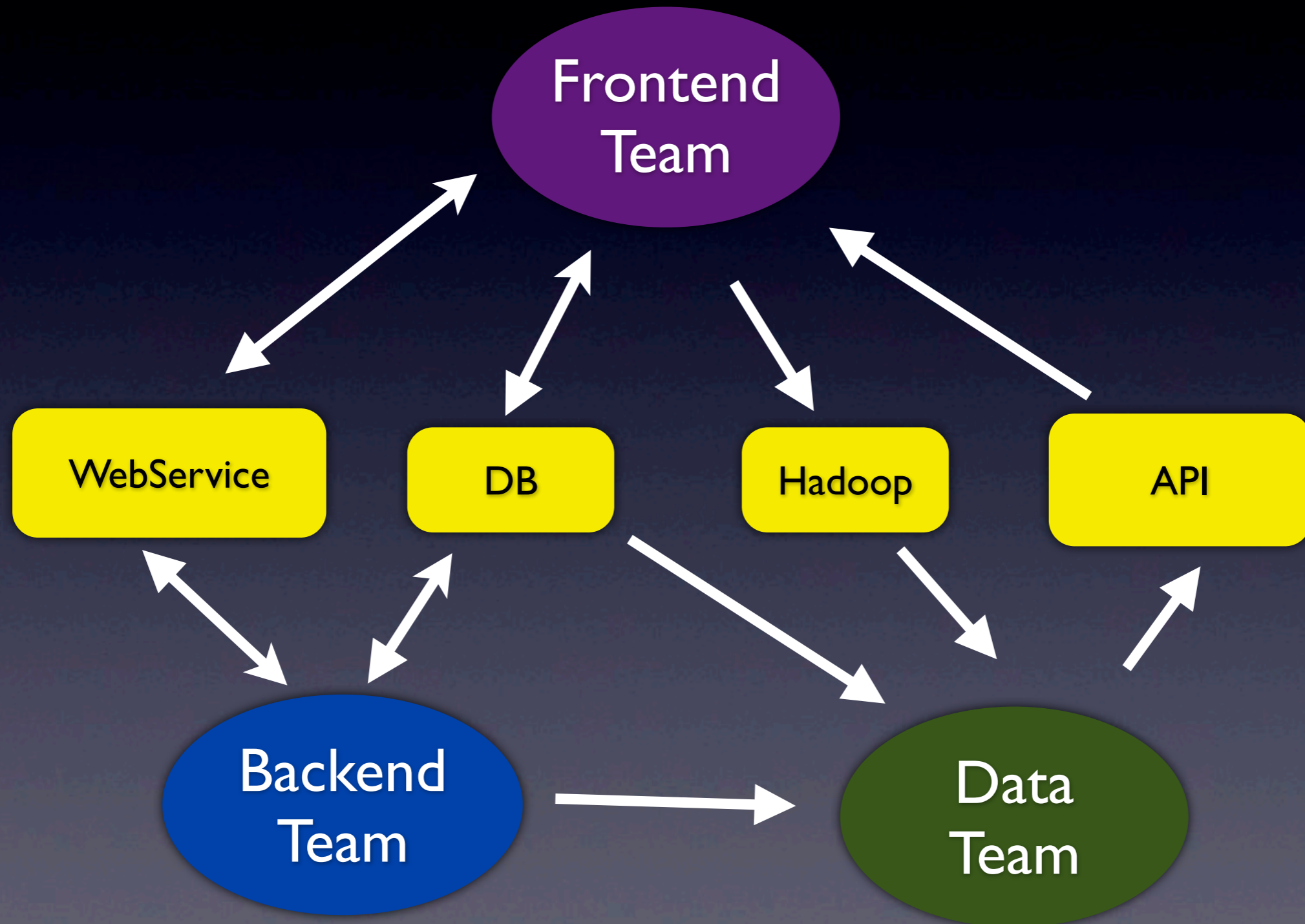
Machine Learning

Data Mart

Business Intelligence

CRM

Problem



We need a **multi-platform**
build and release solution

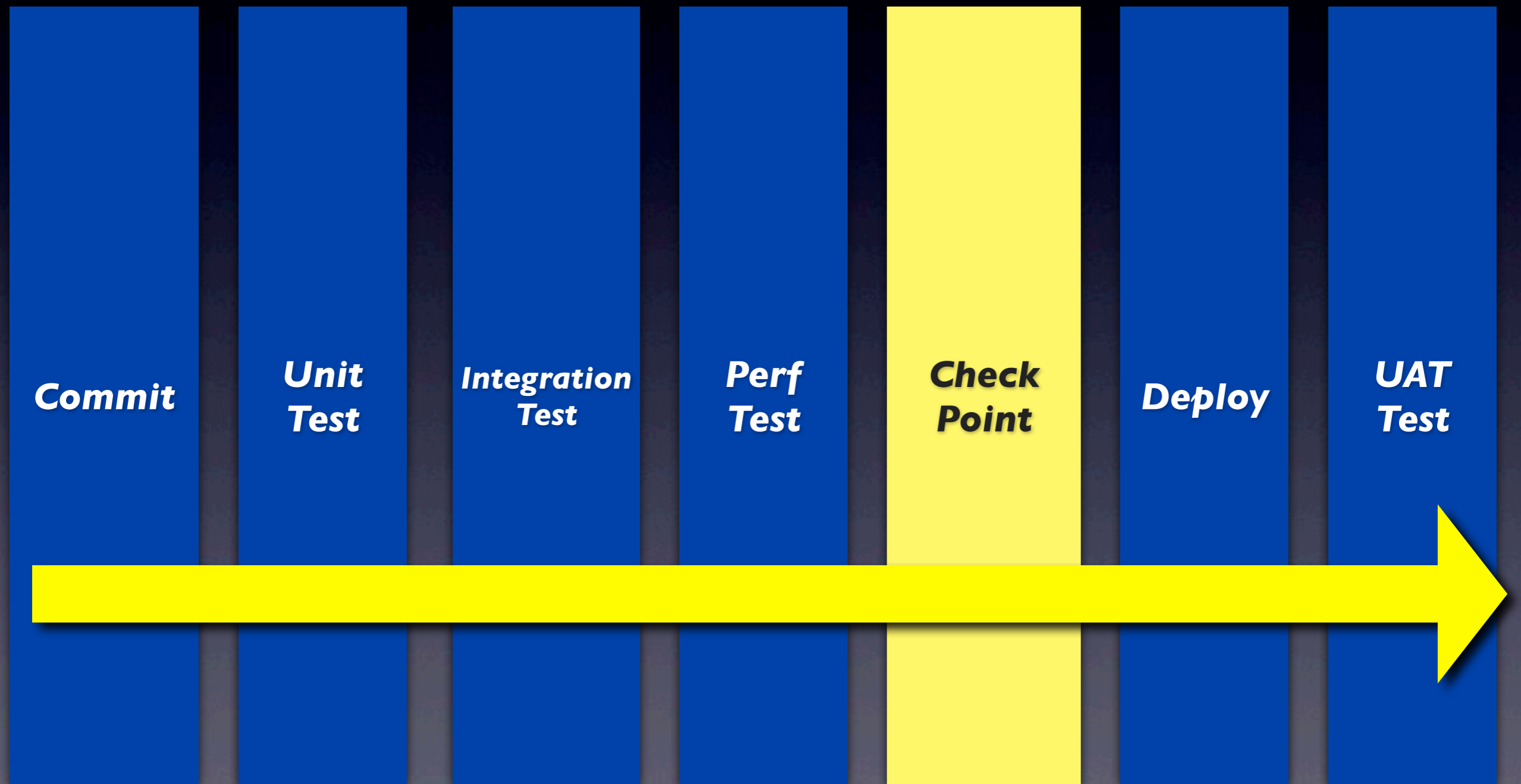
Solution

- Create B&R process
- Build CI infrastructure
- Automate everything

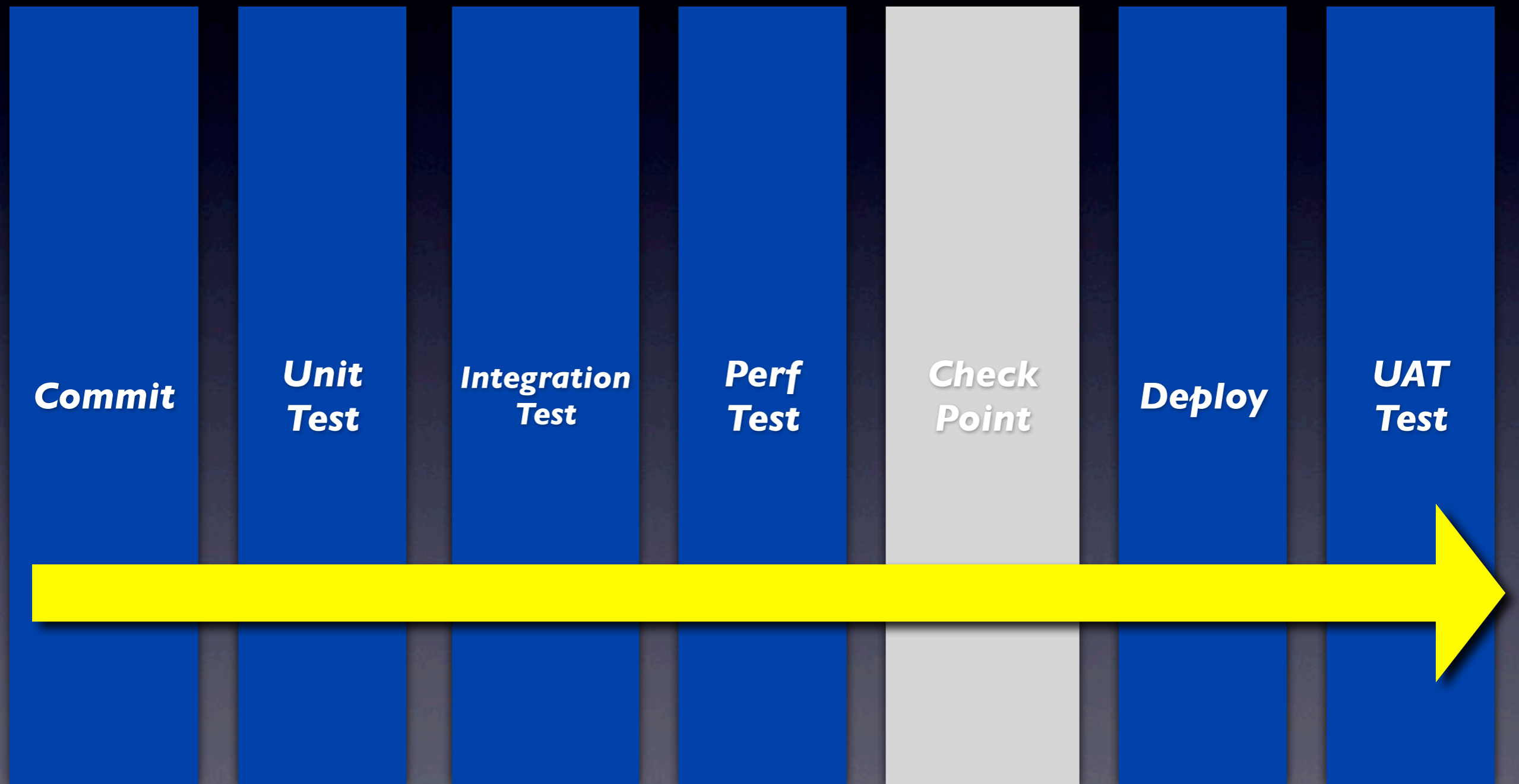
Solution

- Create B&R process
- Build CI infrastructure
- Automate everything

Build and Release Process



Hadoop Build and Release Process



Original Difference

	Frontend Team	Backend Team	Data Team
Process	Scrum	Waterfall	Waterfall
OS	Linux	Windows	Linux
Language	PHP/JAVA	.NET	Pig/MapReduce
Issue Tracking	Bugzilla	Custom System	N/A
Repos	Subversion	VSS	N/A
CI	N/A	N/A	N/A
Int Test	N/A	N/A	N/A
Release Management	Routine RE	Another RE	SDE

Multiplatform Complexity Reduction

	Frontend Team	Backend Team	Data Team
Process	Scrum(not finish yet)		
OS	Linux	Windows	Linux
Language	PHP/JAVA	.NET	Pig/MapReduce
Issue Tracking	Bugzilla		
Repos	Subversion(migrate to git)		
CI	Jenkins		
Int Test	Selenium		
Release Management	Delicated Release Manager		

Communication Cost

The image shows the Japanese character 巨 (Gyo) in a bright yellow color. It is a vertical character with a thick left stroke and three horizontal strokes on the right, resembling a ladder or a staircase.

Huge

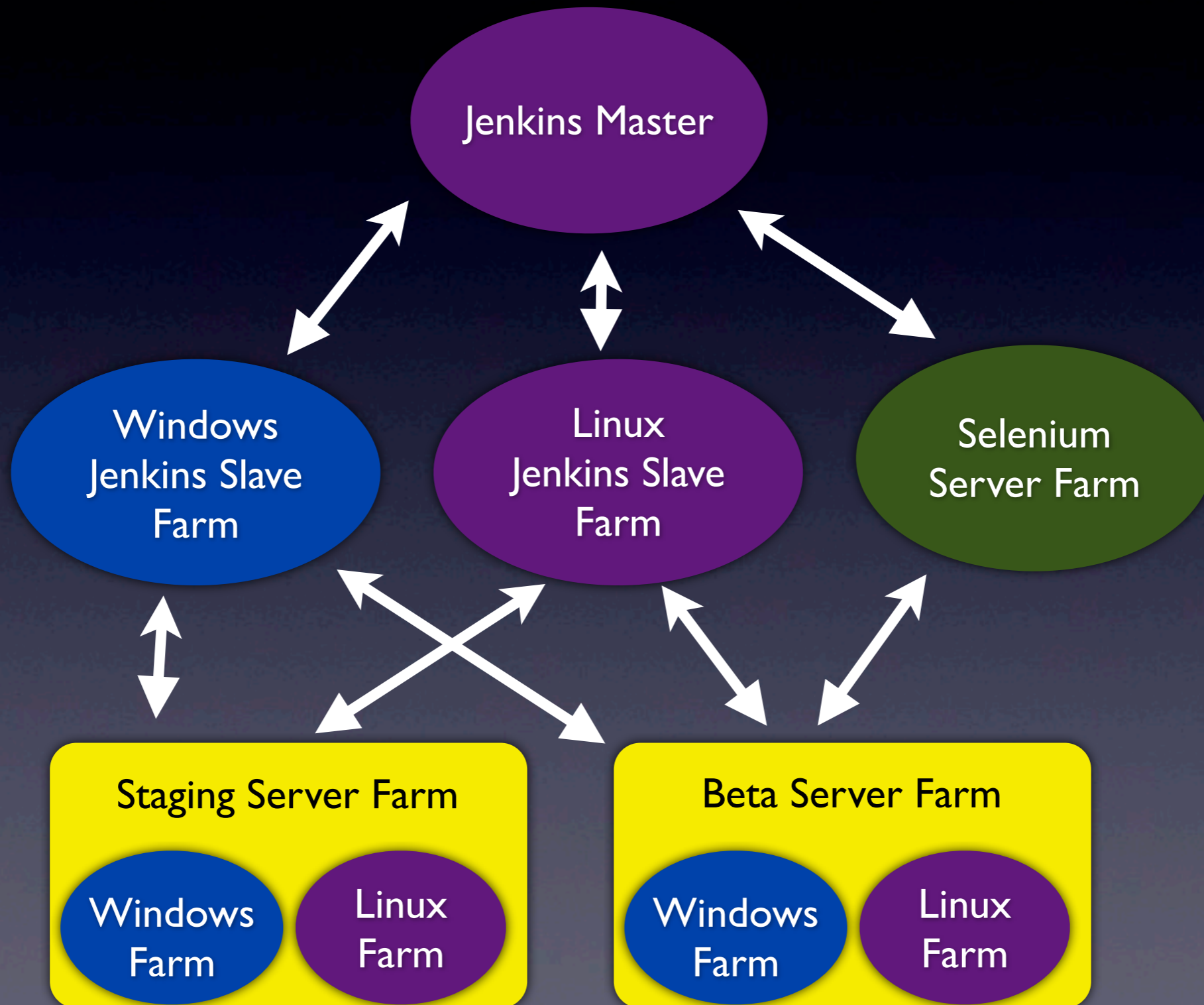
The image shows the Japanese character 大 (Dai) in a bright pink color. It is a vertical character with a thick left stroke and two curved strokes on the right, resembling a large 'A' or a wide 'V'.

Large

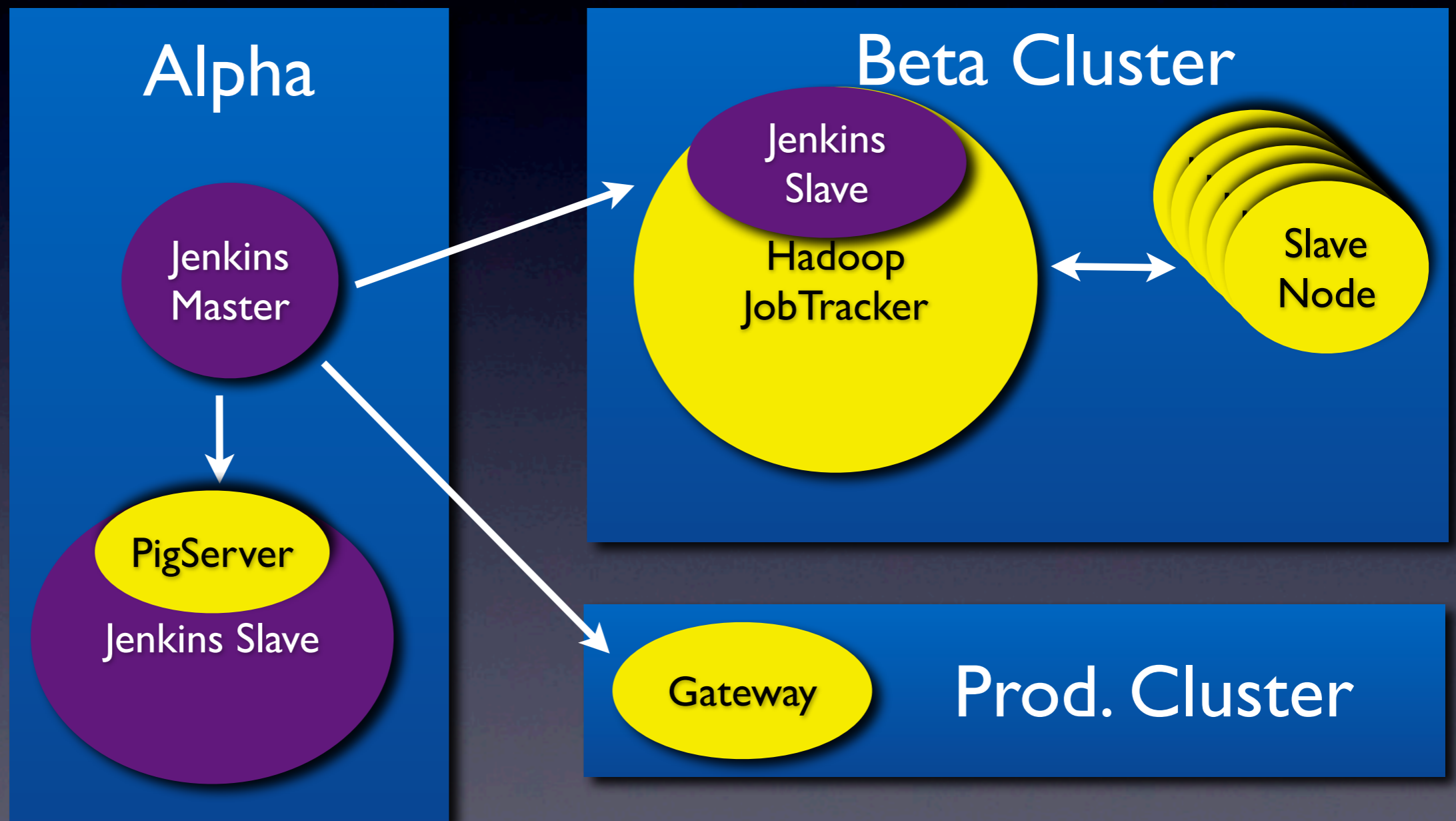
Solution

- Create B&R process
- **Build CI infrastructure**
- Automate everything

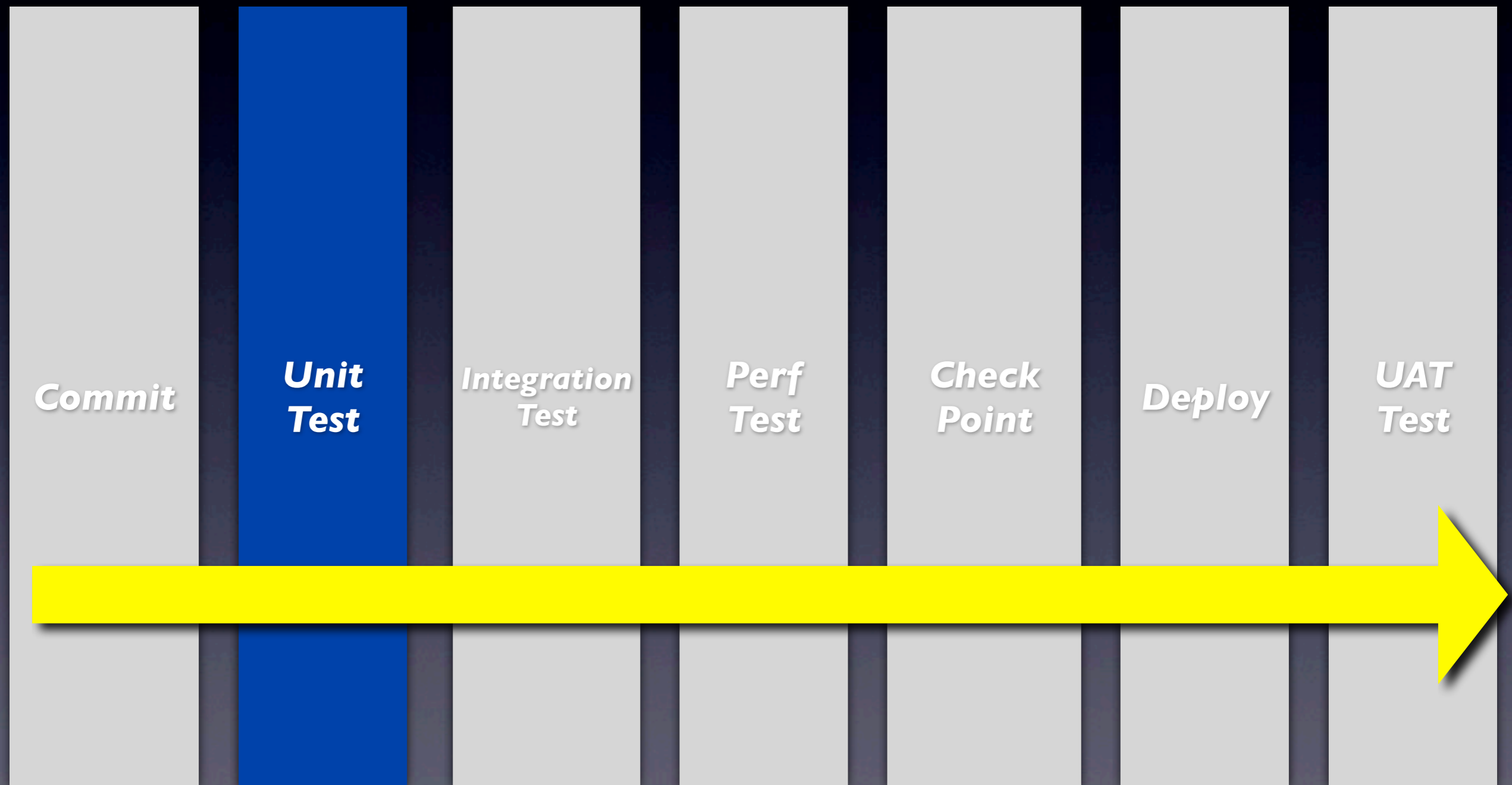
WebSite CI Infrastructure



Hadoop CI Infrastructure



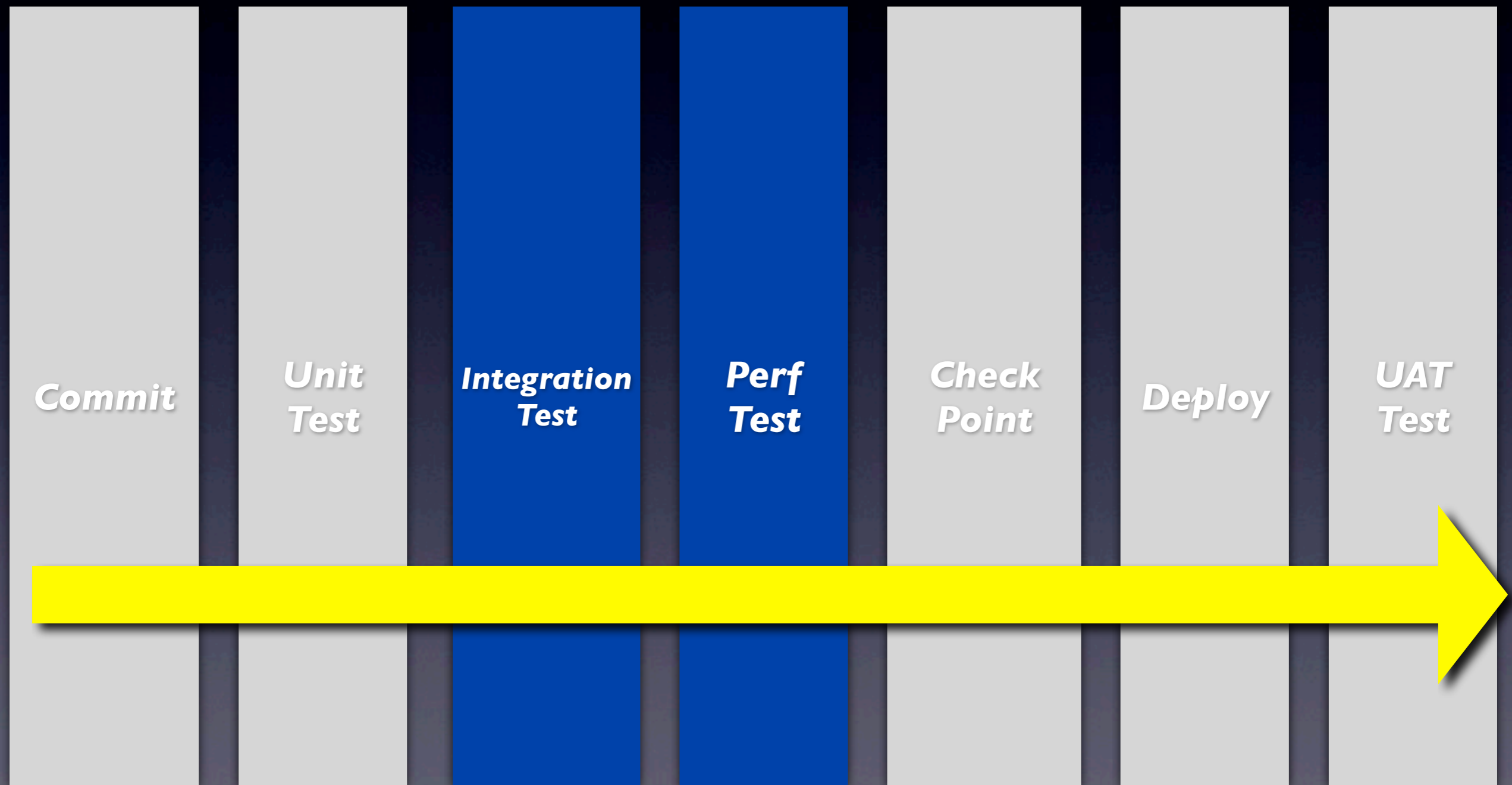
Build and Release Process



PigUnit

- A simple xUnit framework
- No cluster set up is required in local mode
- We change default PigUnit behavior
 - All PigUnit tests run on single pig server
 - Speed up PigUnit

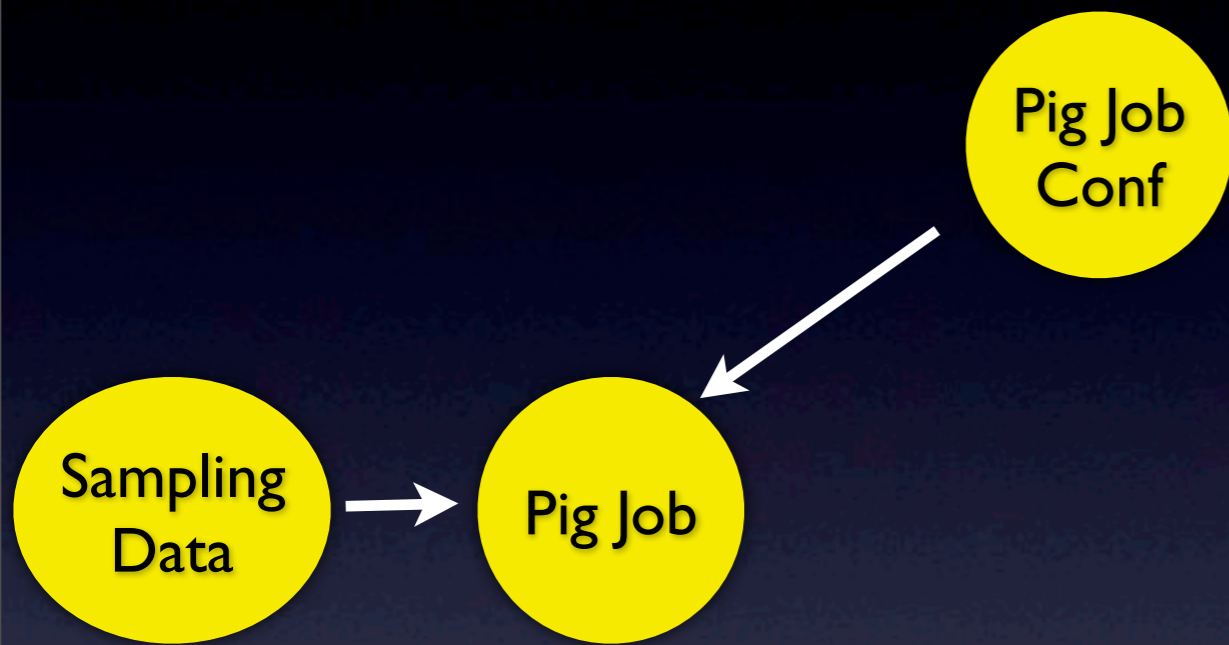
Build and Release Process



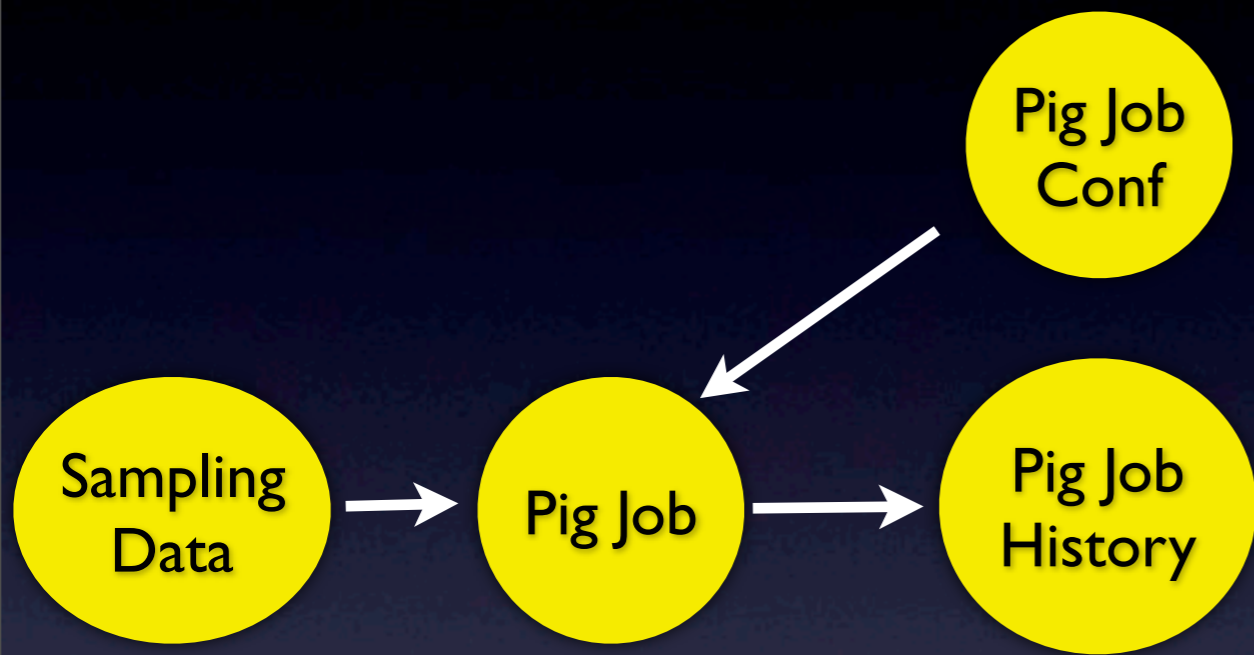
Vaidya

- Rule based performance diagnosis of M/R jobs
- Extendable framework
 - You can add your own rules
 - Write complex rules using existing rules

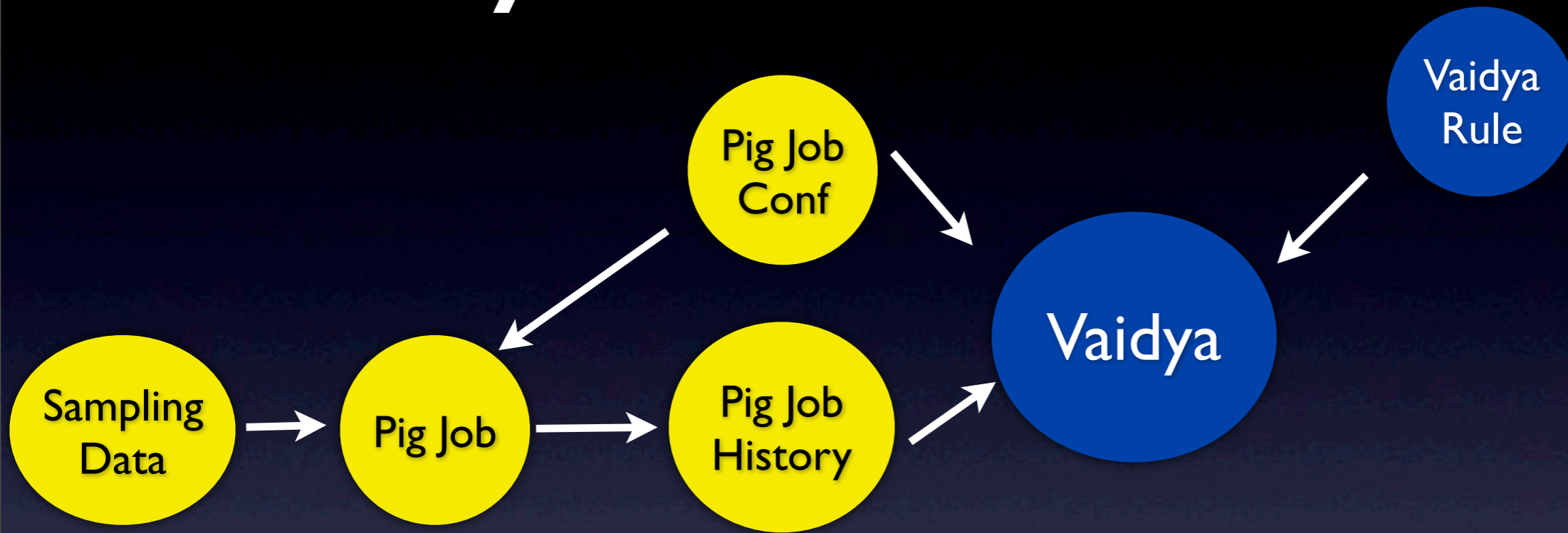
Vaidya Performance Test



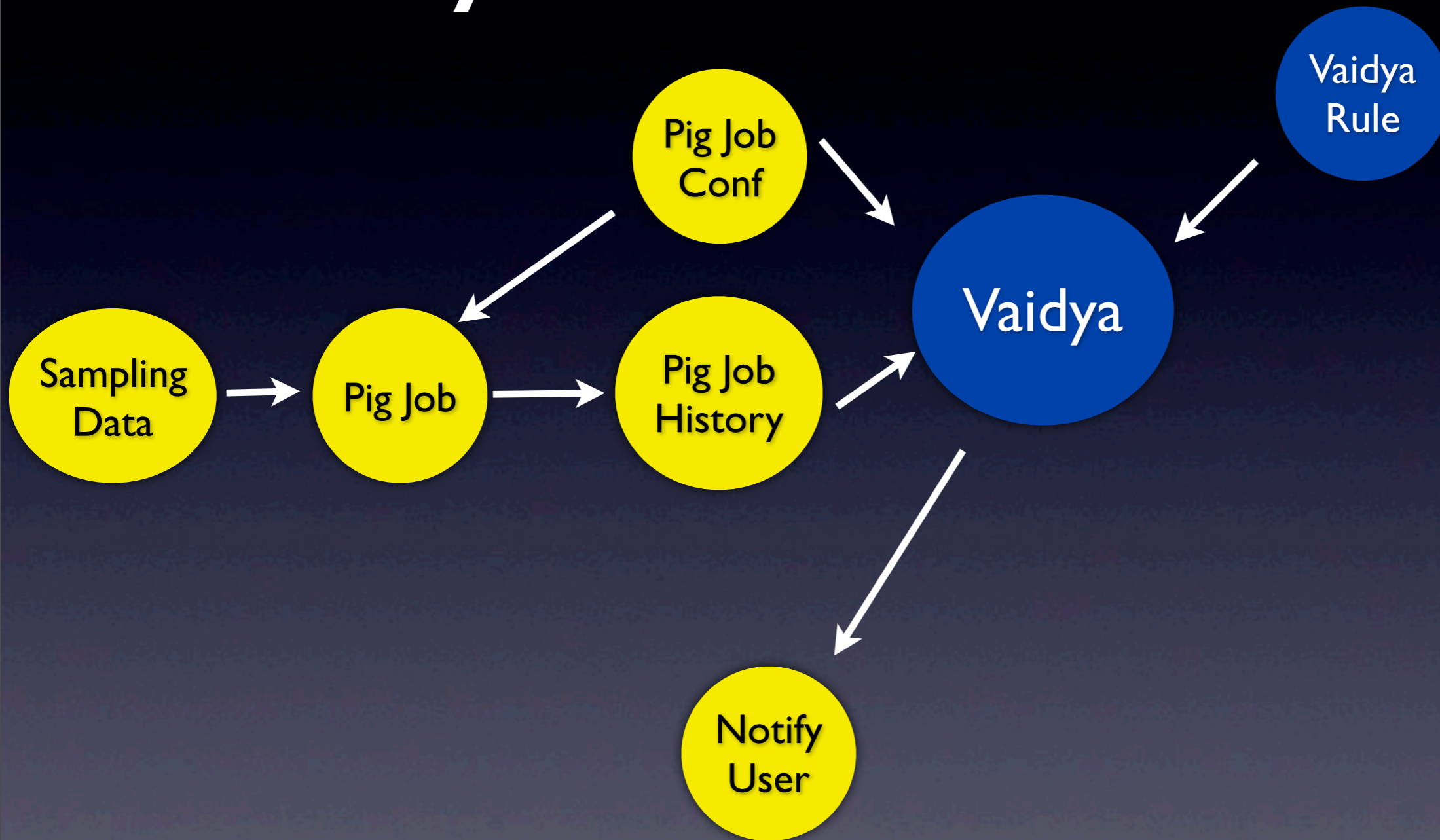
Vaidya Performance Test



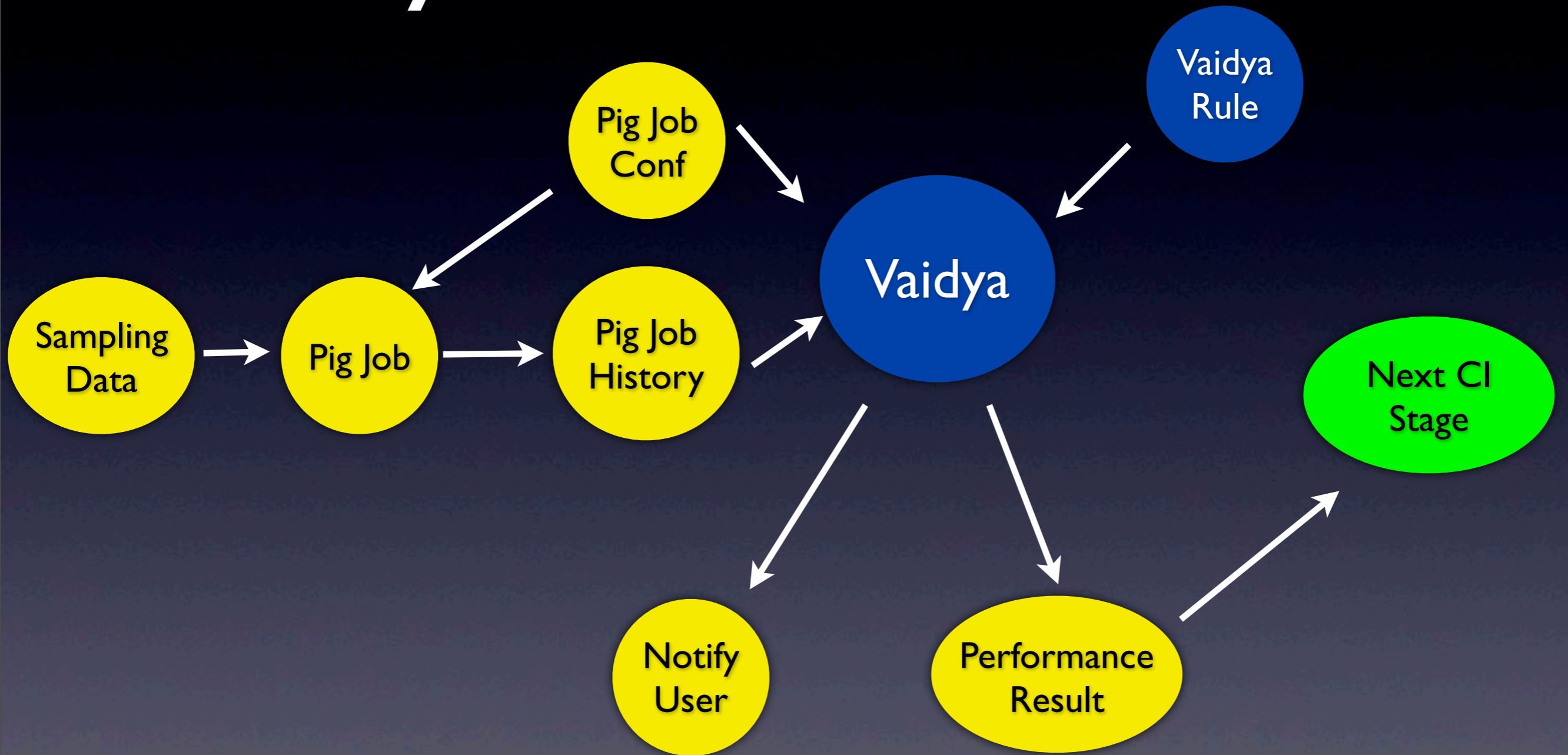
Vaidya Performance Test



Vaidya Performance Test



Vaidya Performance Test



Vaidya Rule Example

```
<DiagnosticTest>
```

```
<Title><![CDATA[Balanced Reduce Partitioning]]></Title>
```

```
<ClassName>
```

```
<![CDATA[org.apache.....BalancedReduceP
```

```
</ClassName>
```

```
<Description>
```

```
<![CDATA[This rule tests .....
```

```
</Description>
```

```
<Importance><![CDATA[High]]></Importance>
```

```
<SuccessThreshold><![CDATA[0.20]]></SuccessThreshold>
```

```
<Prescription><![CDATA[advice]]></Prescription>
```

```
</DiagnosticTest>
```

See if the reduce job is balanced or not

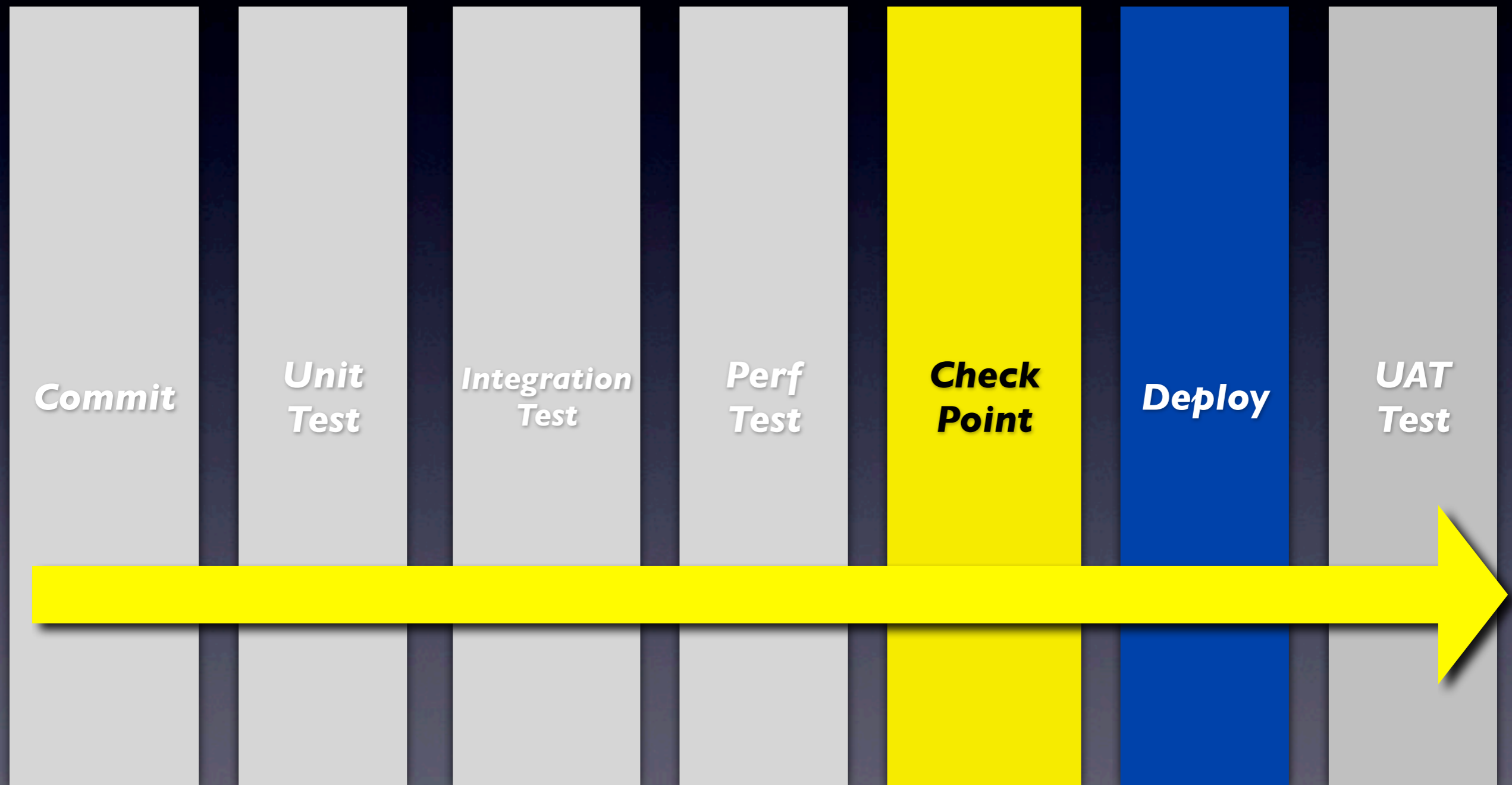
Rule Importance

Diagnosis success threshold

Solution

- Create B&R process
- Build CI infrastructure
- Automate everything

Build and Release Process



Before Deploy

- Check bug status
- Documentation
 - Create a git tag
 - Auto doc generating on bugzilla
- Deploy to production cluster
- Change bug status

Auto comment in bugzilla

Comment 25 by [Wisely Chen \(thegiive\)](#) 26 hours ago at 20:33 PDT

CI push code to production grid done.

The build source repos url is in

https://git.corp.yahoo.com/thegiive/pig/tree/Prod-2013-06-17_08-47-42

The release note is in

https://git.corp.yahoo.com/thegiive/pig/blob/Prod-2013-06-17_08-47-42/release_note.

If you need more information, please check

<http://boiledcoiled.corp.sg3.yahoo.com:8080/jenkins/job/GenReleaseNoteAnd>

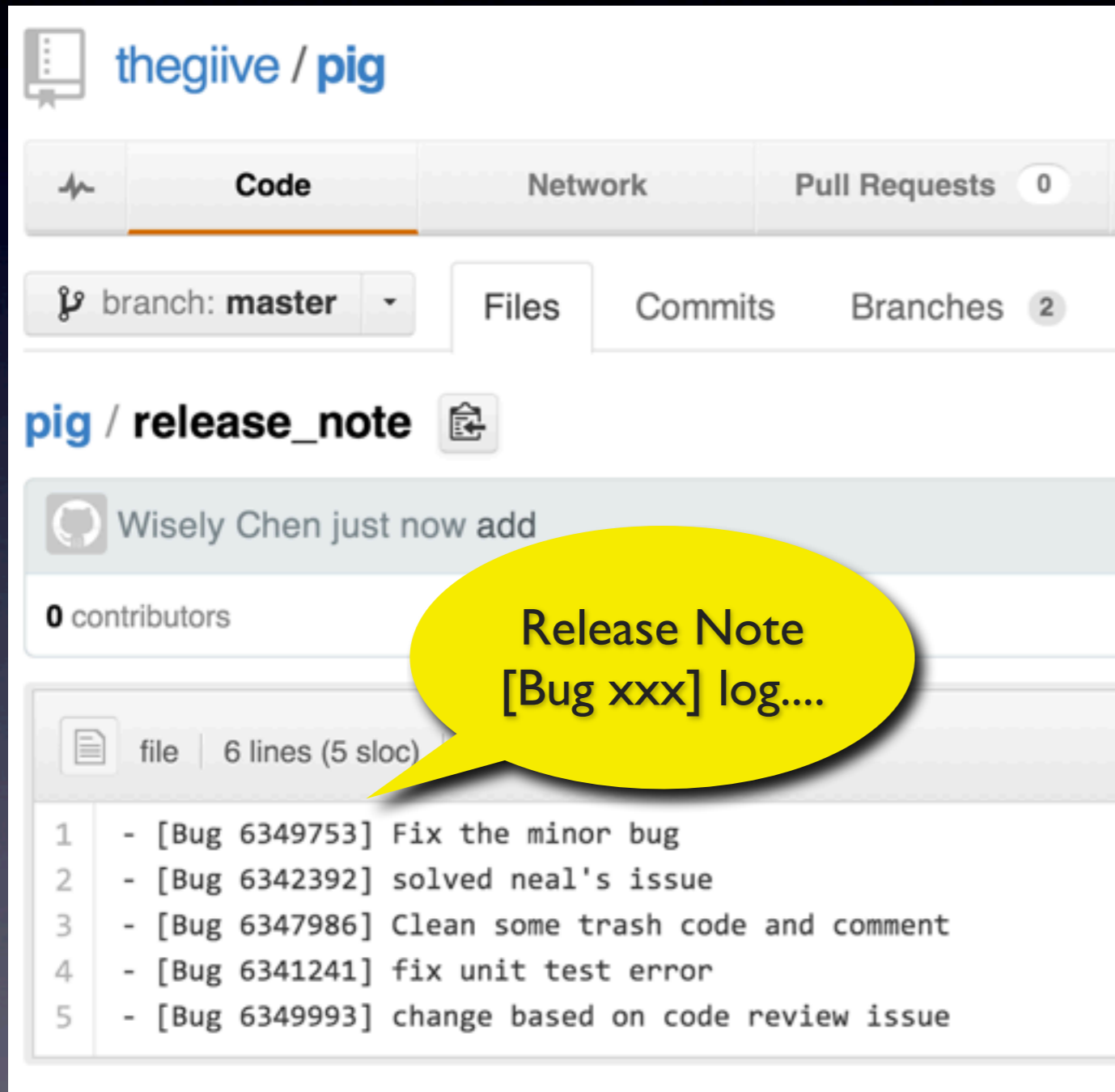
Fields Changed	From	To
Status	REOPENED	RESOLVED
Target Version	Unspecified	inProd
Resolution	---	FIXED

Repos url

Release Note

Issue status change

Auto create git tag



thegiive / pig

Code Network Pull Requests 0

branch: master Files Commits Branches 2

pig / release_note

Wisely Chen just now add

0 contributors

file | 6 lines (5 sloc)

```
1 - [Bug 6349753] Fix the minor bug
2 - [Bug 6342392] solved neal's issue
3 - [Bug 6347986] Clean some trash code and comment
4 - [Bug 6341241] fix unit test error
5 - [Bug 6349993] change based on code review issue
```

Release Note
[Bug xxx] log....

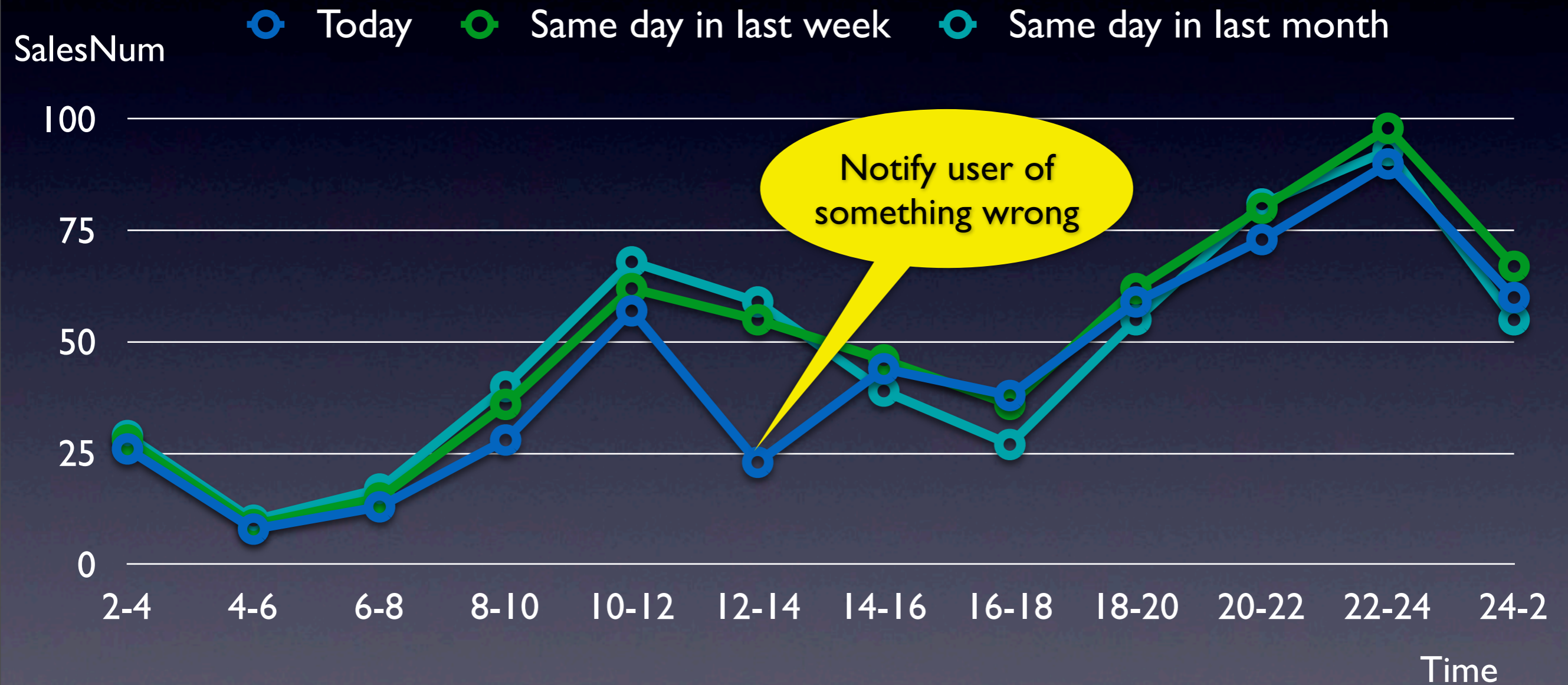


Git Tag

Tags

a day ago	Prod-2013-06-17_08-47-42 - Jenkins Git plugin tagging 504d470 zip tar.gz
a day ago	Prod-2013-06-17_08-39-06 - Jenkins Git plugin tagging be57791 zip tar.gz
a day ago	Prod-2013-06-17_08-29-28 - Jenkins Git plugin tagging 3b15c79 zip tar.gz
a day ago	Prod-2013-06-17_08-23-24 - Jenkins Git plugin tagging 2ed51be zip tar.gz
a day ago	Prod-2013-06-17_08-17-54 - Jenkins Git plugin tagging c7b5c13 zip tar.gz
2 days ago	Prod-2013-06-16_02-26-04 - Jenkins Git plugin tagging c6c7ebf zip tar.gz

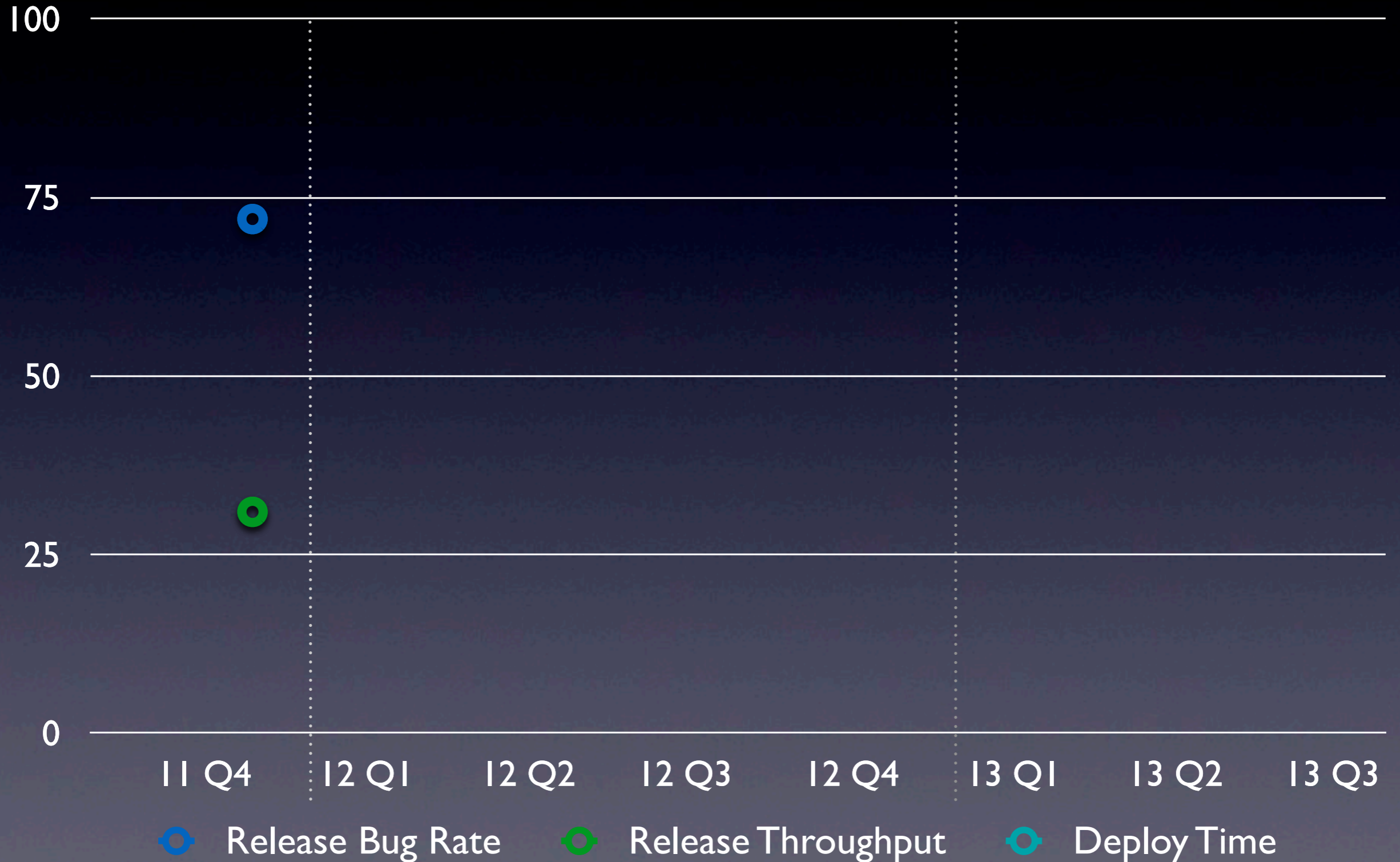
Application Logic Monitor



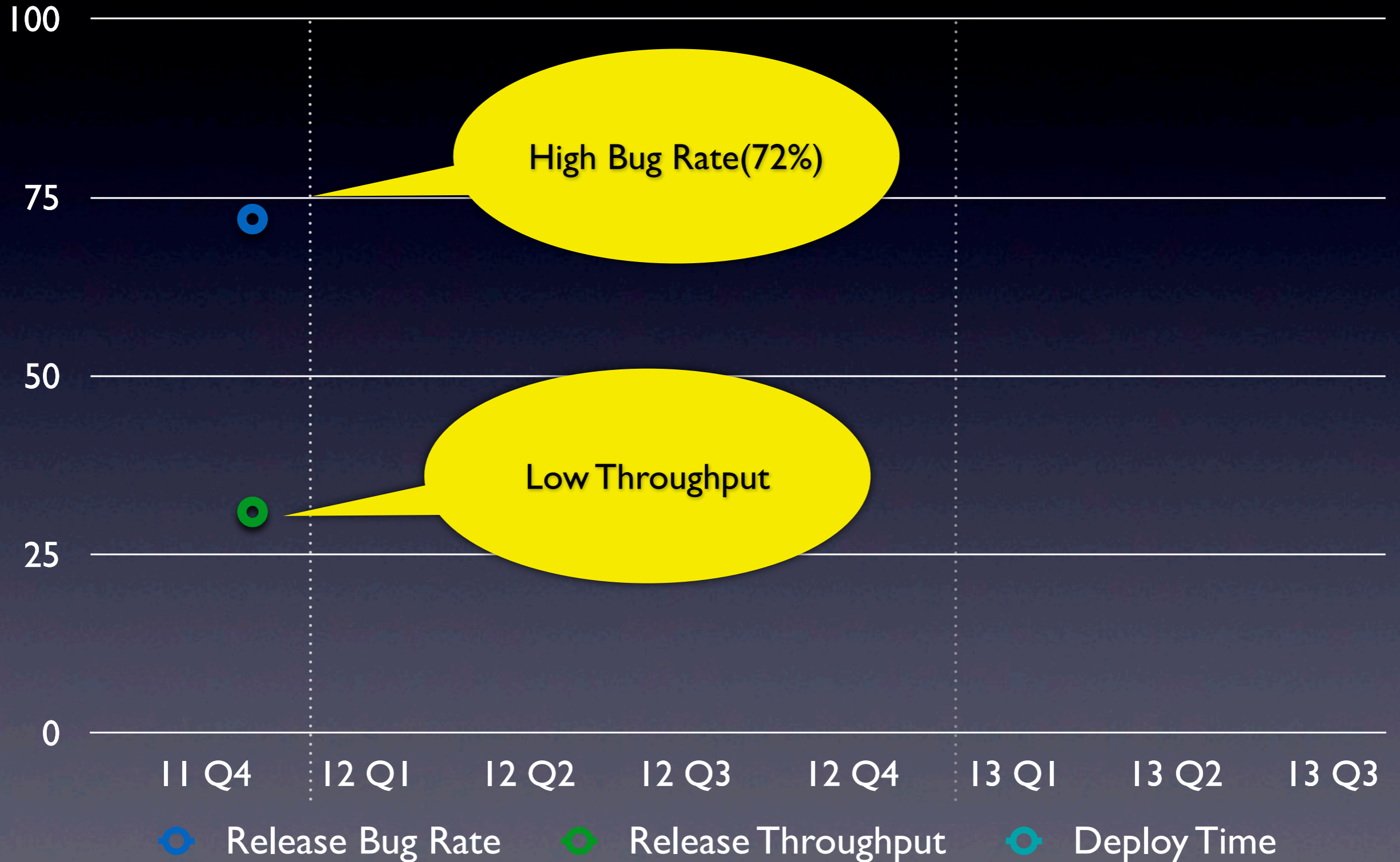
Monitor today's sales revenue with baseline

Result

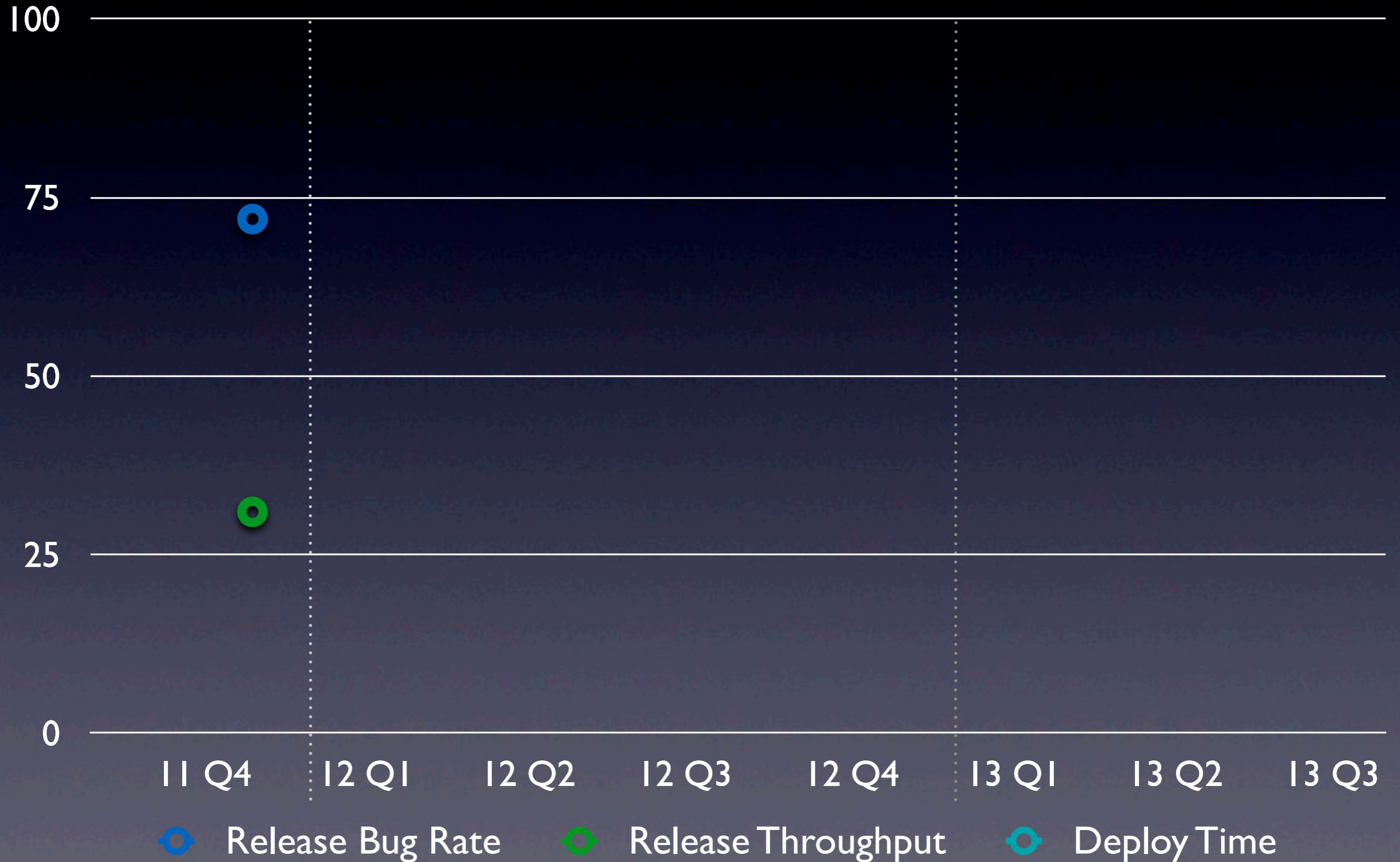
11 Q4: Original



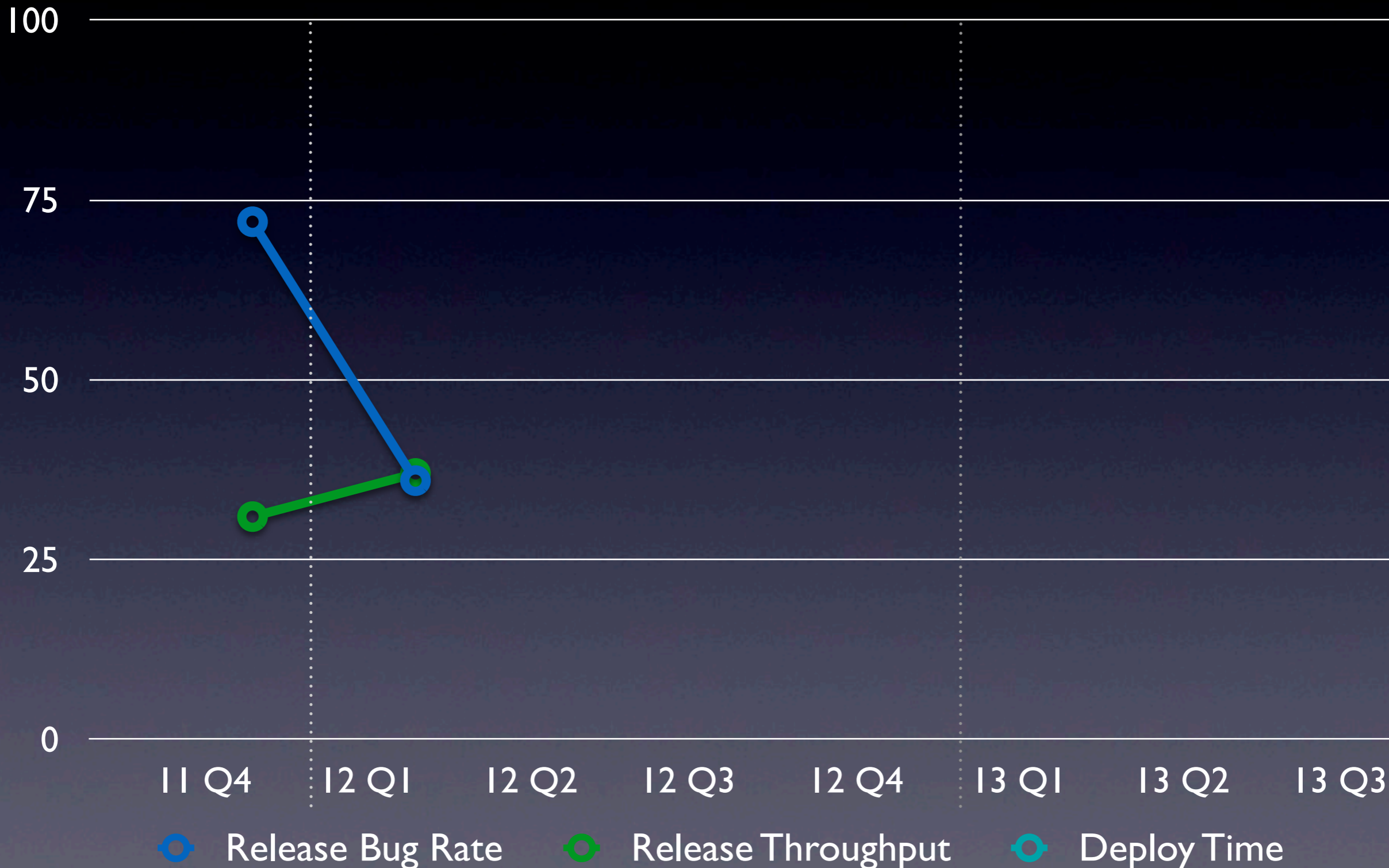
11 Q4: Original



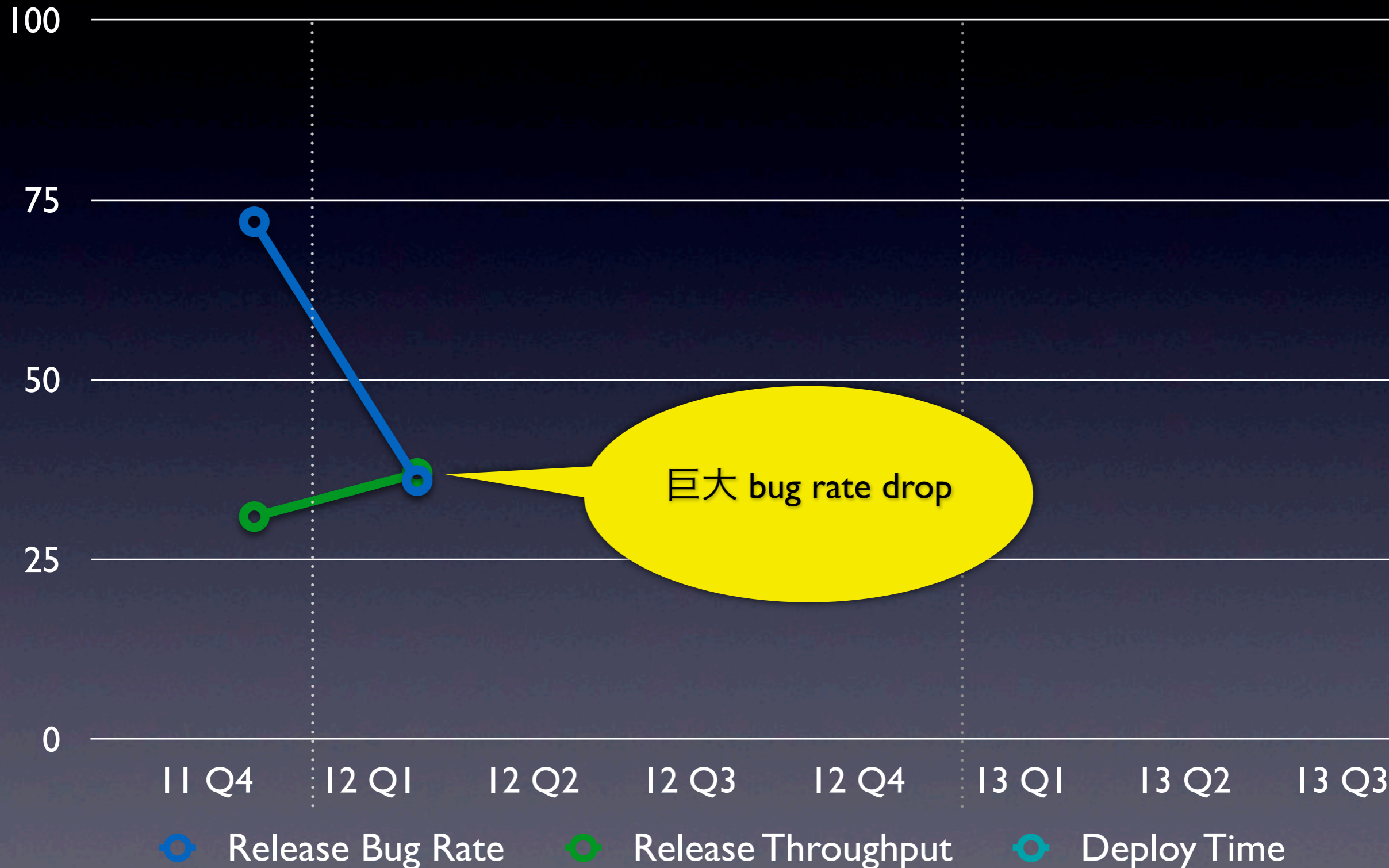
11 Q4: Original



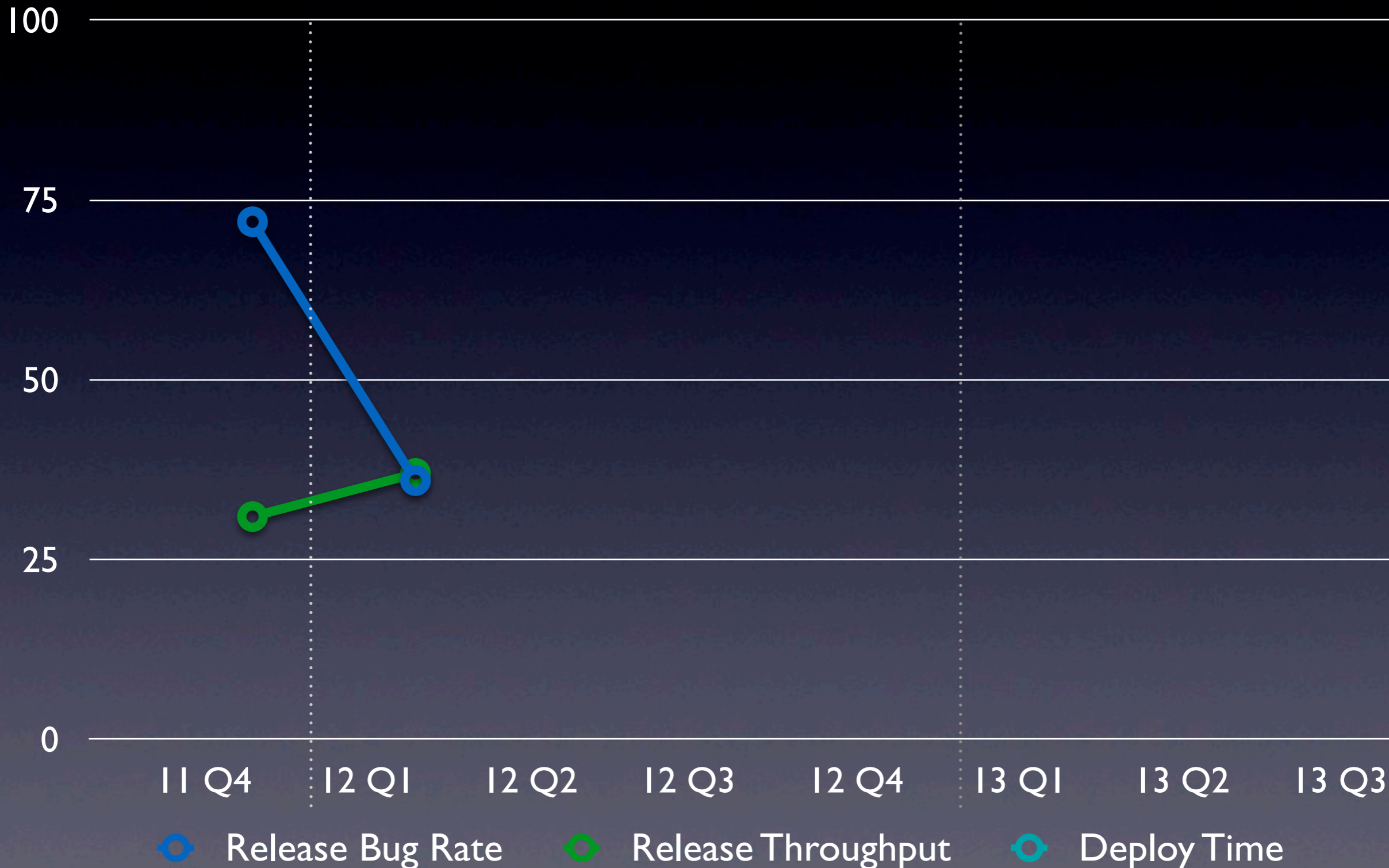
I2Q1 : Create B&R Process...



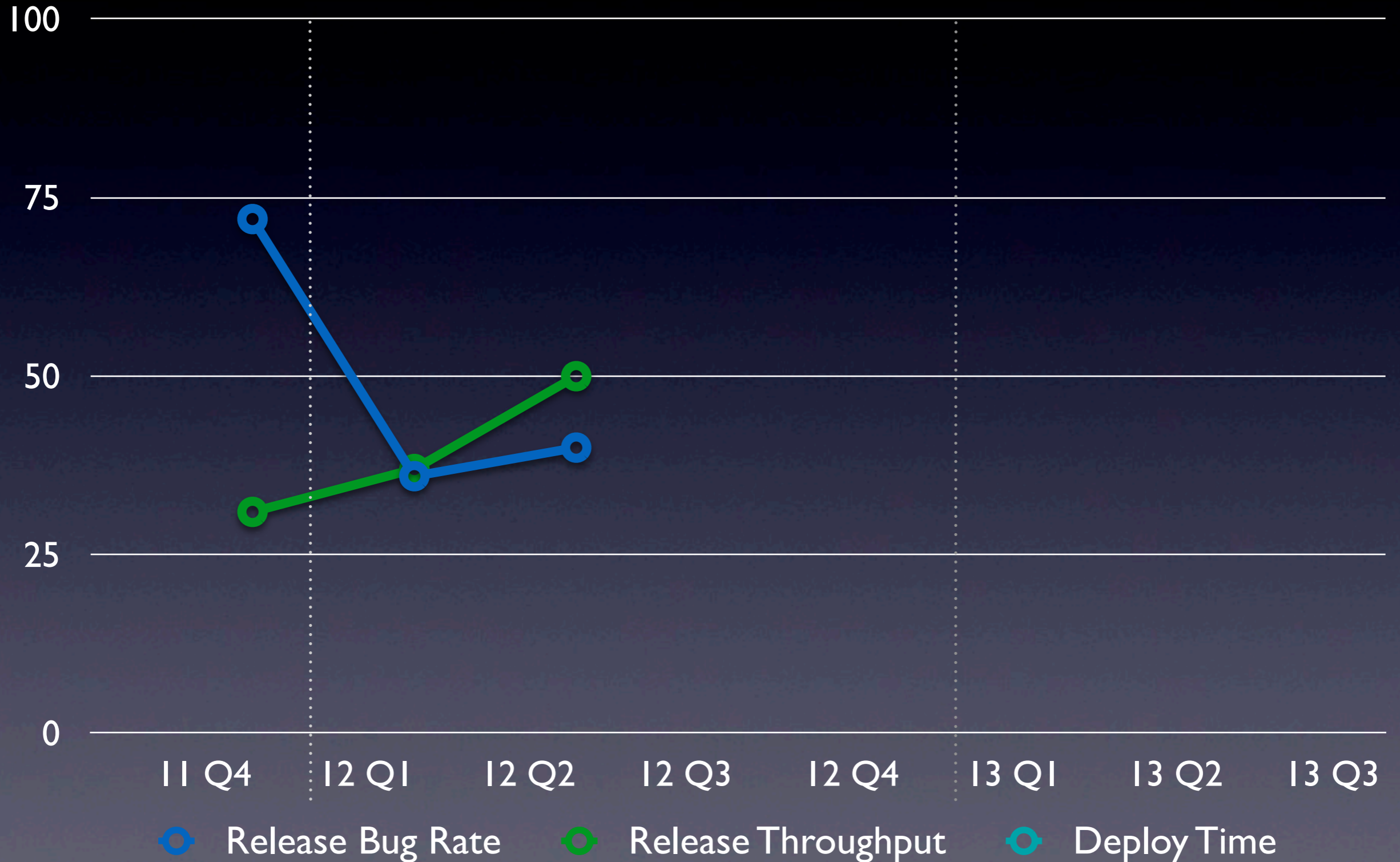
I2Q1 : Create B&R Process...



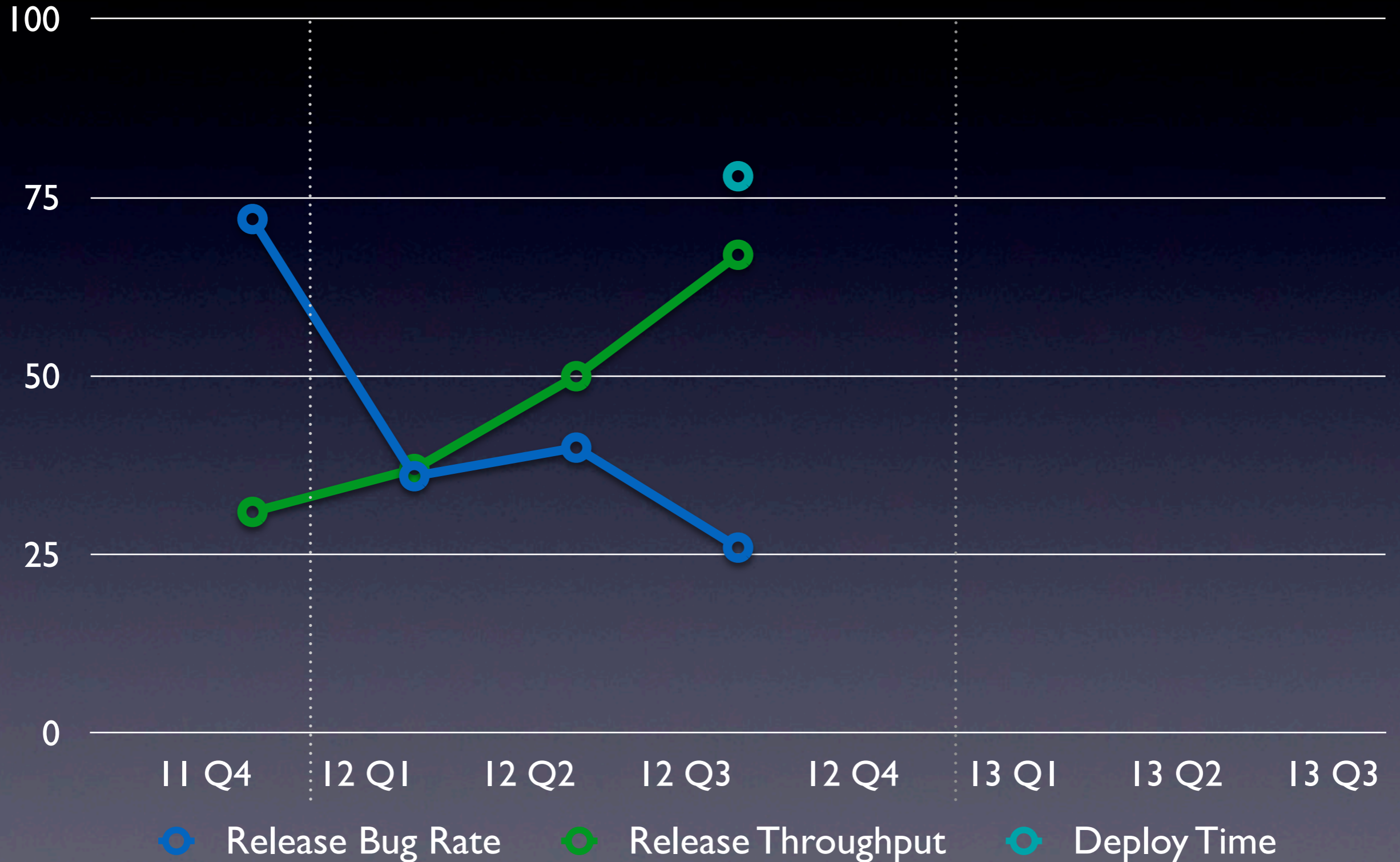
I2Q1 : Create B&R Process...



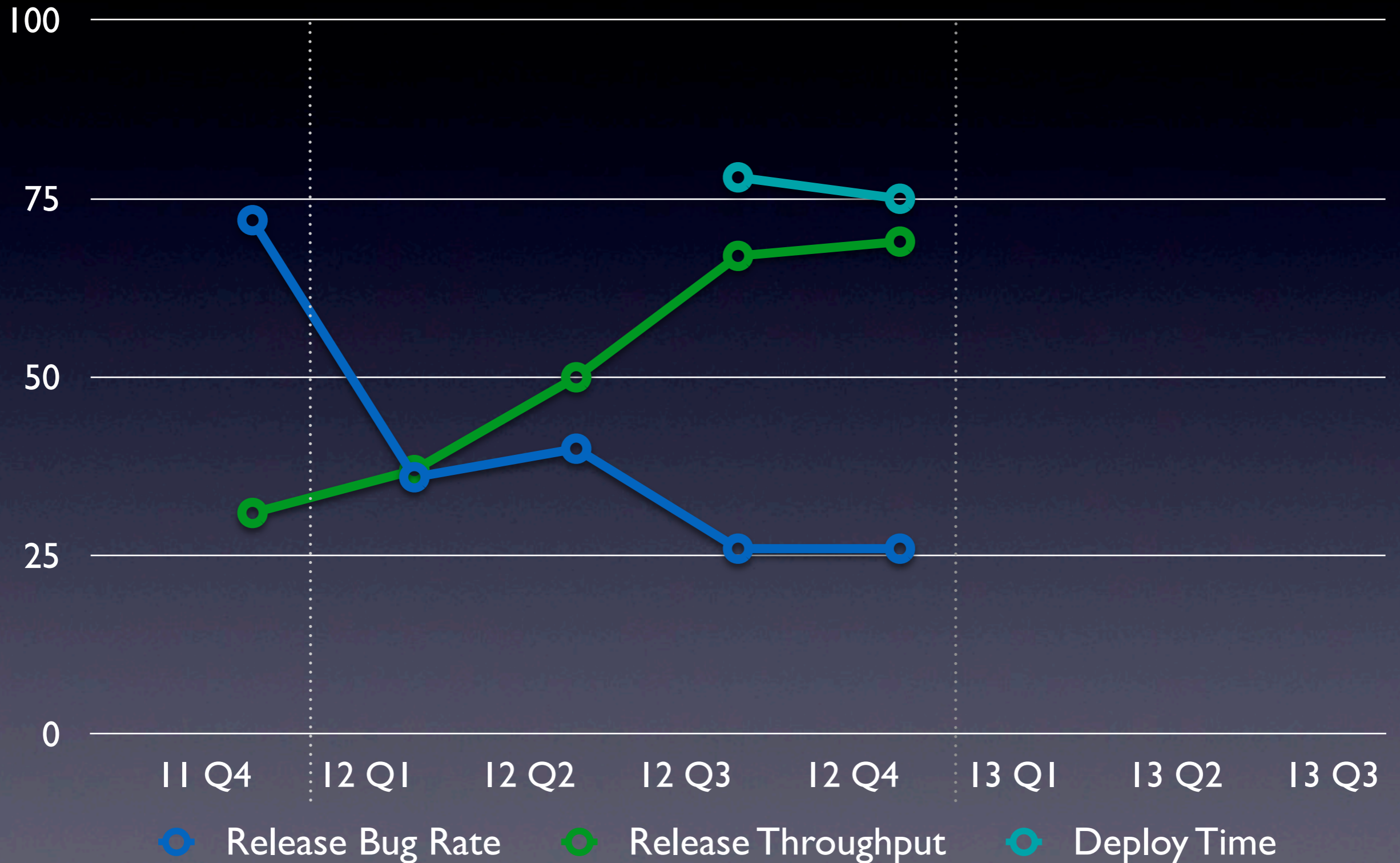
I2Q2 : Build CI Infrastructure...



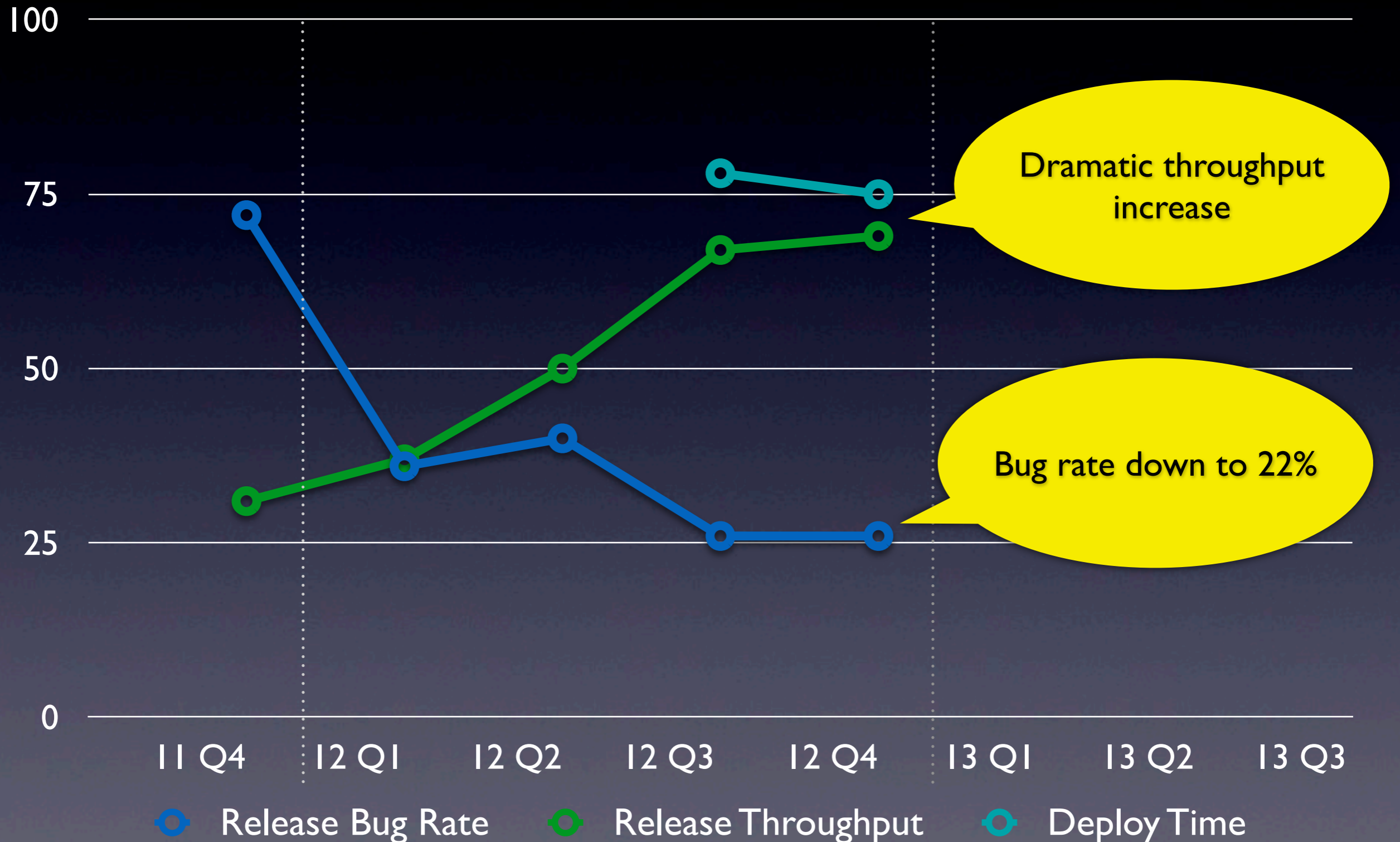
I2Q3 : Build CI Infrastructure...



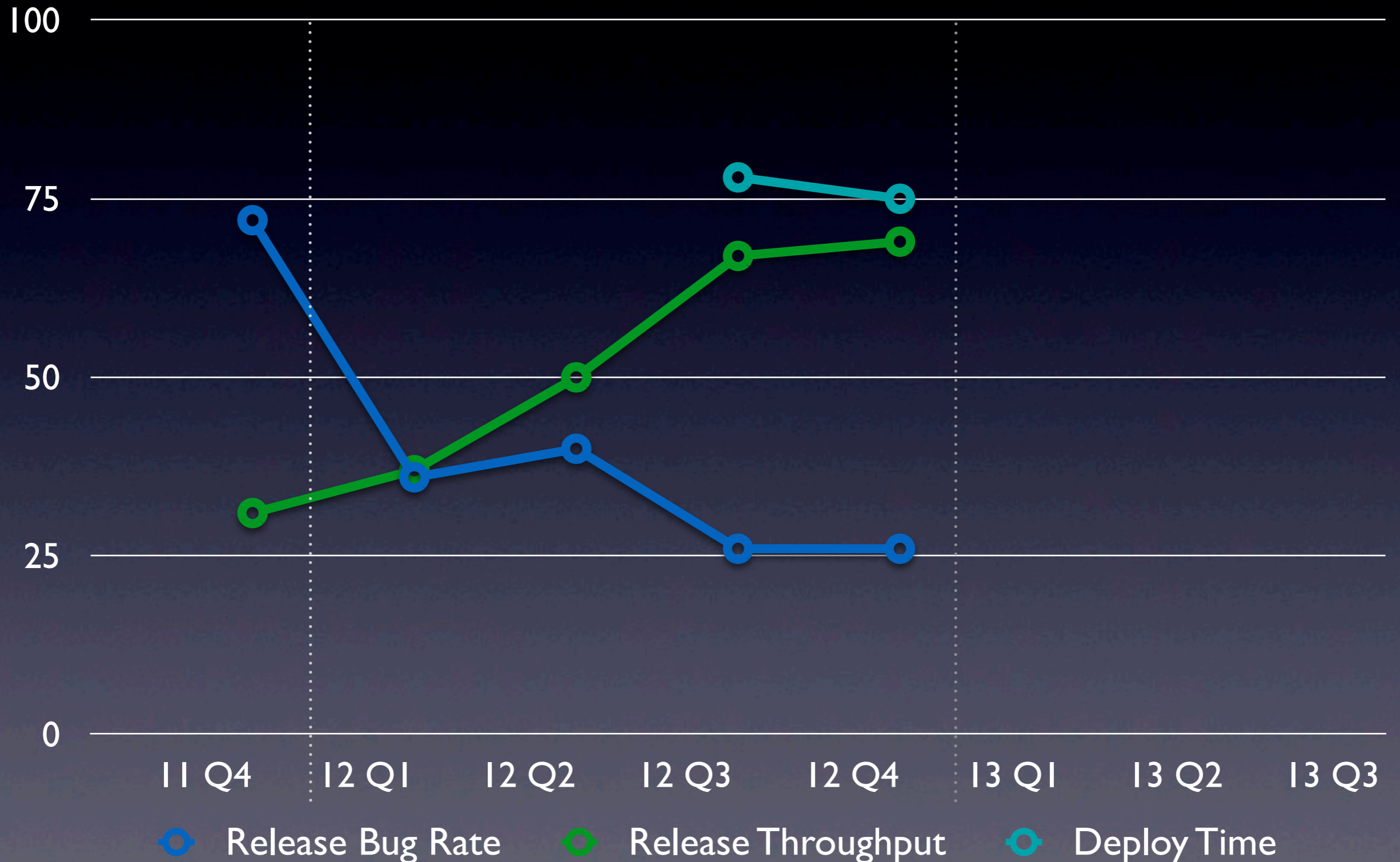
I2Q4 : Build CI Infrastructure...



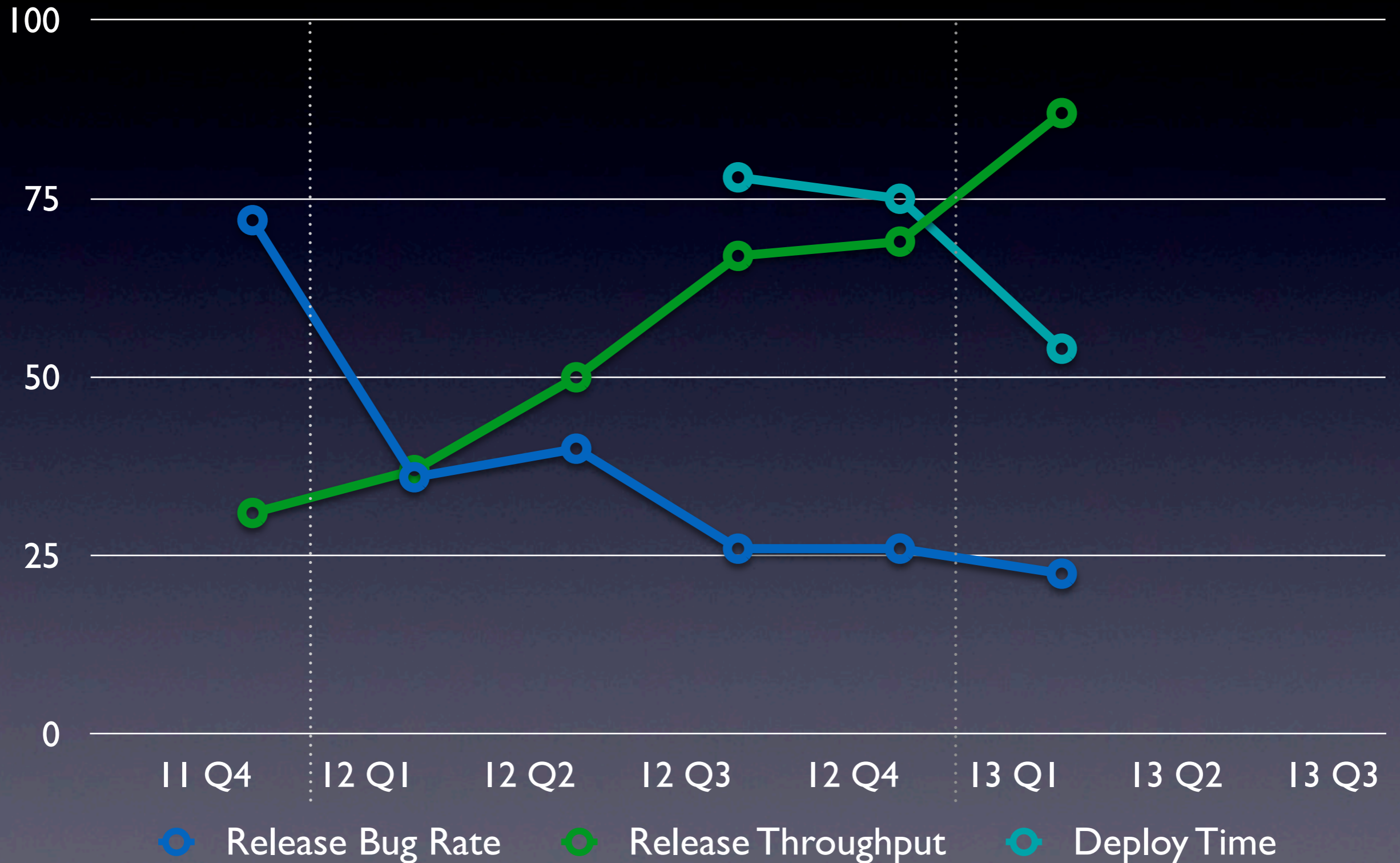
I2Q4 : Build CI Infrastructure...



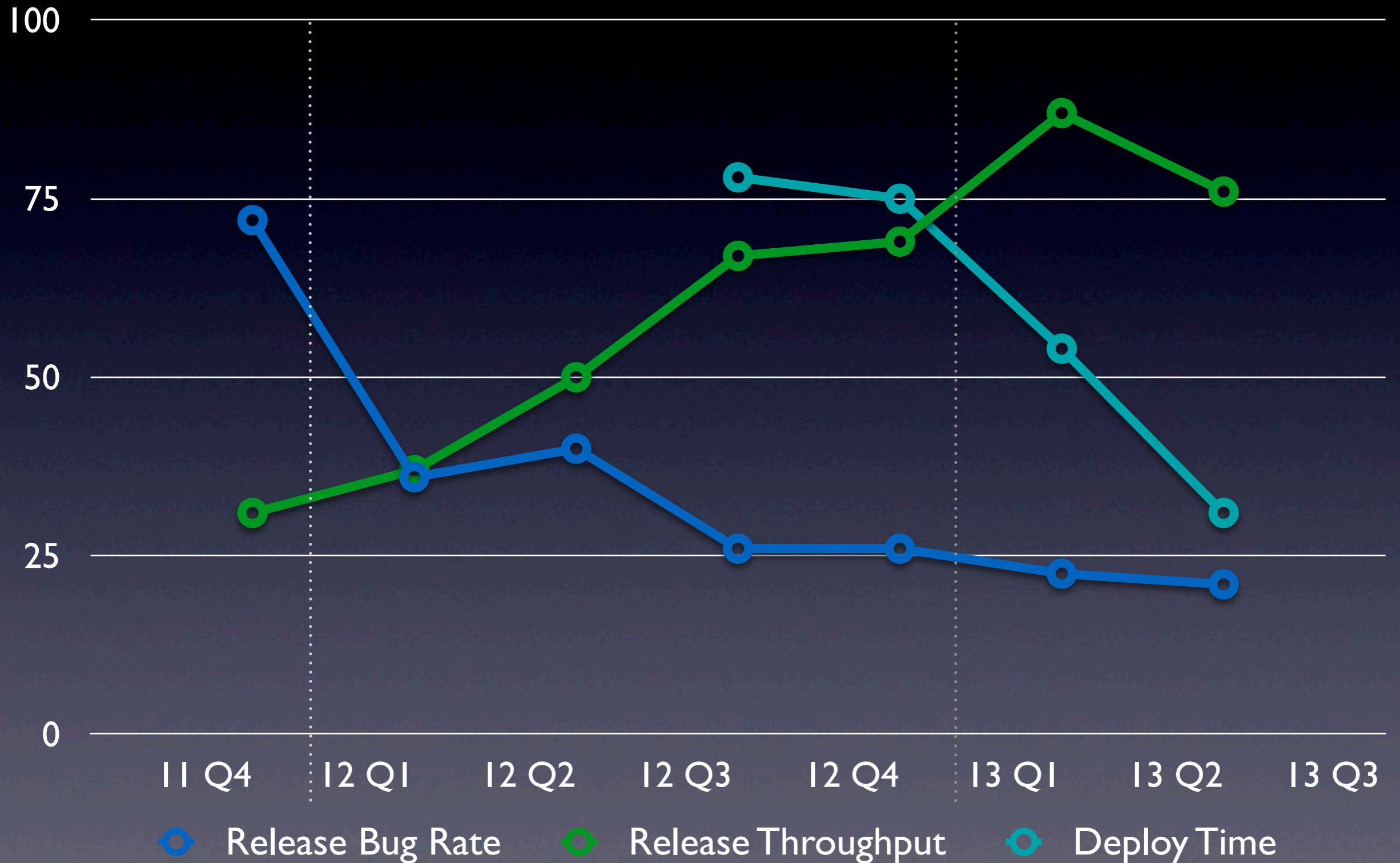
I2Q4 : Build CI Infrastructure...



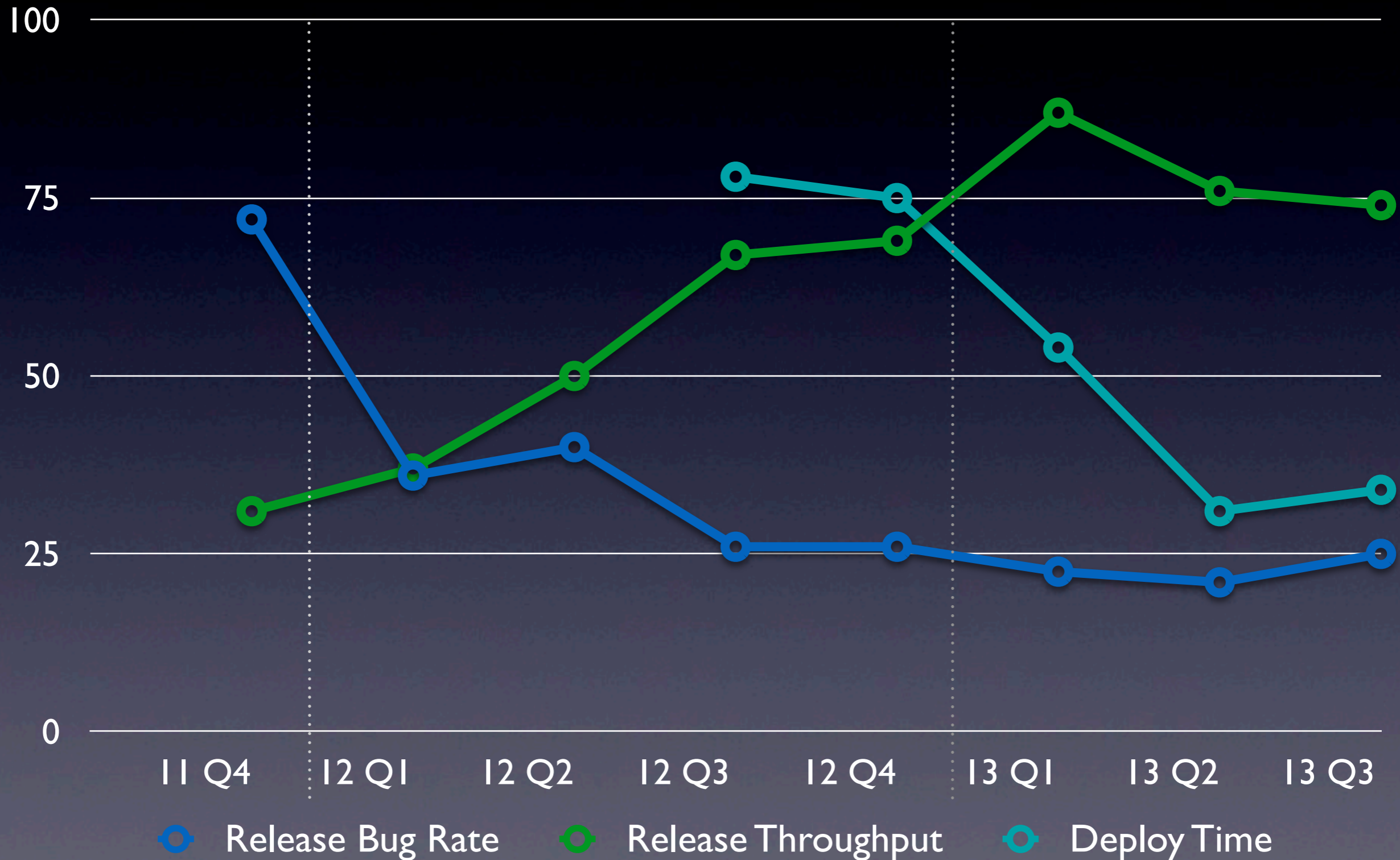
I3Q1 : Automate Everything...



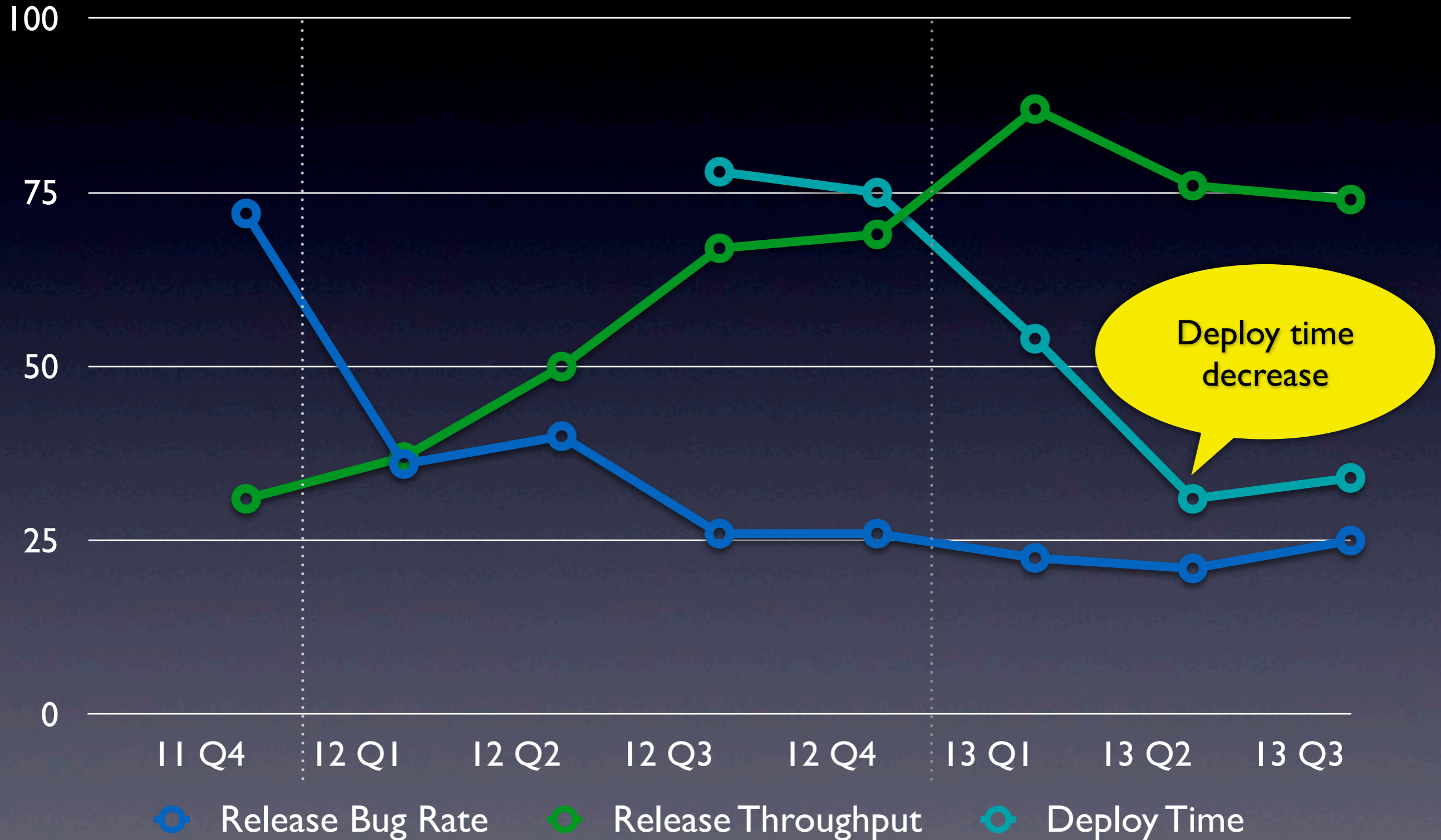
I3Q2 : Automate Everything...



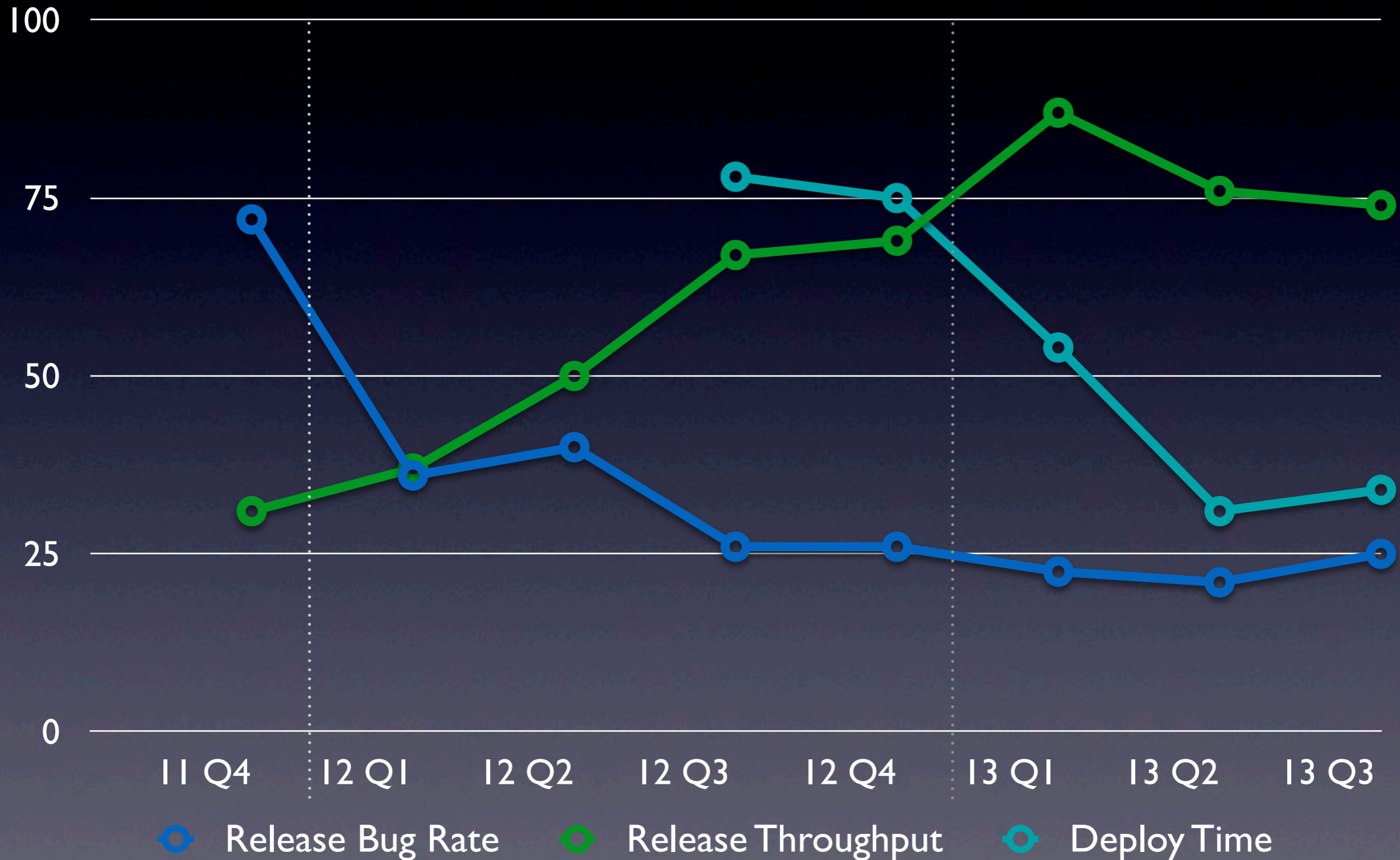
I 3Q3: Automate Everything...



I 3Q3: Automate Everything...



I 3Q3: Automate Everything...



Hadoop CI Demo

- Demo1 : Unit test
- Demo2 : Performance test
- Demo3 : Deploy

QuickTime Player File Edit View Window Help

boiledcoiled.corp.sg3.yahoo.com:8080/jenkins/view/Hadoop/

Jenkins

Jenkins > Hadoop > ENABLE AUTO REFRESH

Build Pipeline

Run History Configure Add Step Delete Manage

Pipeline	Step 1	Step 2	Step 3	Step 4	Step 5
804	#804 UnitTest Jun 17, 2013 8:37:32 AM 46 sec	#383 Performance... Jun 17, 2013 8:38:23 AM 37 sec	#614 GenReleaseN... Jun 17, 2013 8:39:06 AM 18 sec	#505 ExecGridJob Jun 17, 2013 8:39:29 AM 47 sec	#261 Notify Jun 17, 2013 8:40:22 AM 0.18 sec
803	#803 UnitTest Jun 17, 2013 8:36:21 AM 47 sec	PerformanceTest	GenReleaseNoteAn...	ExecGridJob	Notify
802	#802 UnitTest Jun 17, 2013 8:27:52 AM 47 sec	#382 Performance... Jun 17, 2013 8:28:45 AM 37 sec	#613 GenReleaseN... Jun 17, 2013 8:29:28 AM 18 sec	#504 ExecGridJob Jun 17, 2013 8:29:52 AM 1 min 24 sec	#260 Notify Jun 17, 2013 8:31:21 AM 0.22 sec
801	#801 UnitTest Jun 17, 2013 8:26:41 AM 48 sec	PerformanceTest	GenReleaseNoteAn...	ExecGridJob	Notify

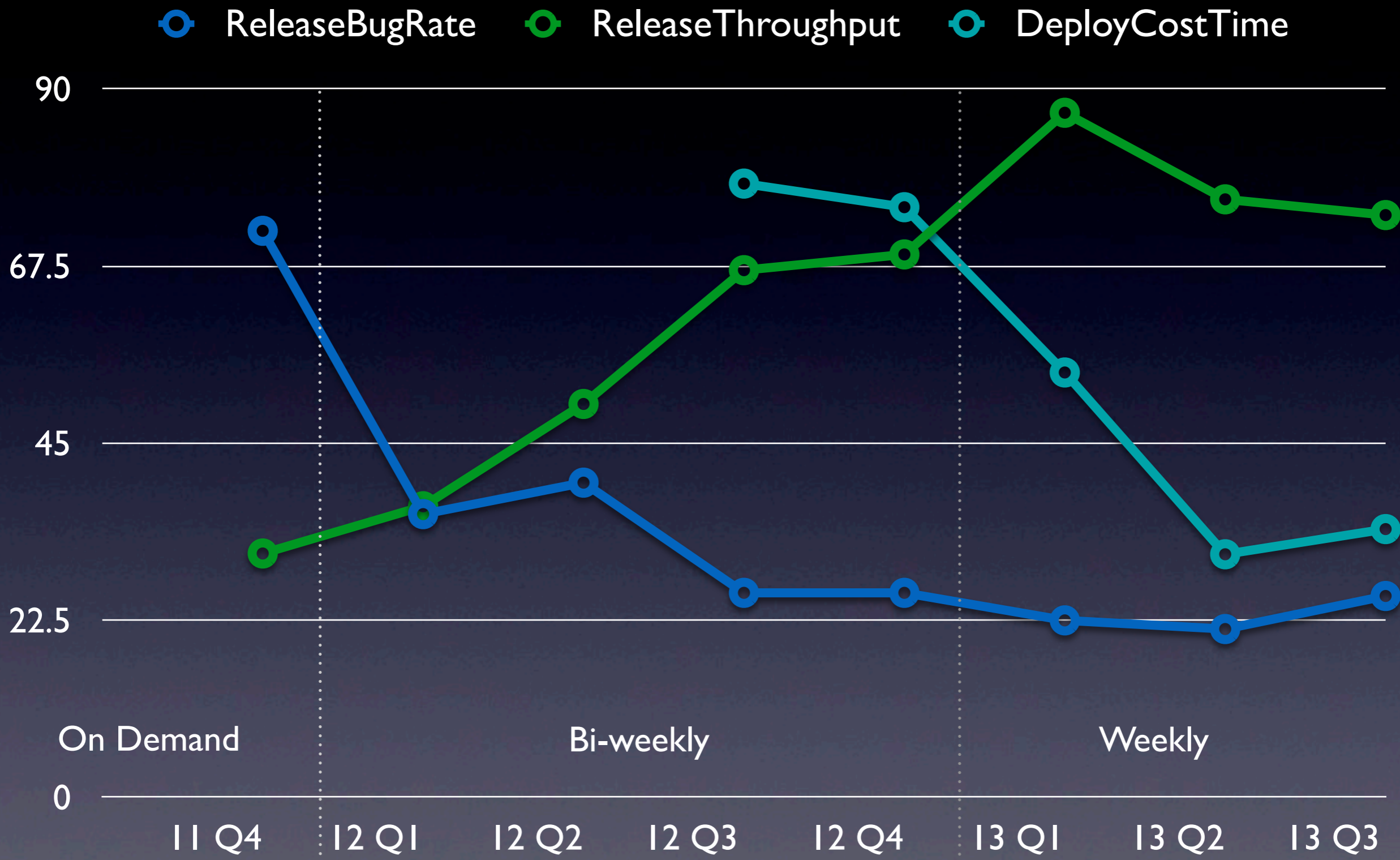
Conclusion

- This process can reduce bug rate
- This process can increase productivity
- Jenkins can fit in any platform

YAHOO!
奇摩

謝謝大家, 祝你有美好的一天

Appendix



Web Site CI Process

Frontend
2 Team
(Linux)

Catalog

Payment

Member

Mobile

Backend
3 Team
(Window)

Cash Flow

Supplier

Logistic

CMS

Data
2 Team
(Hadoop)

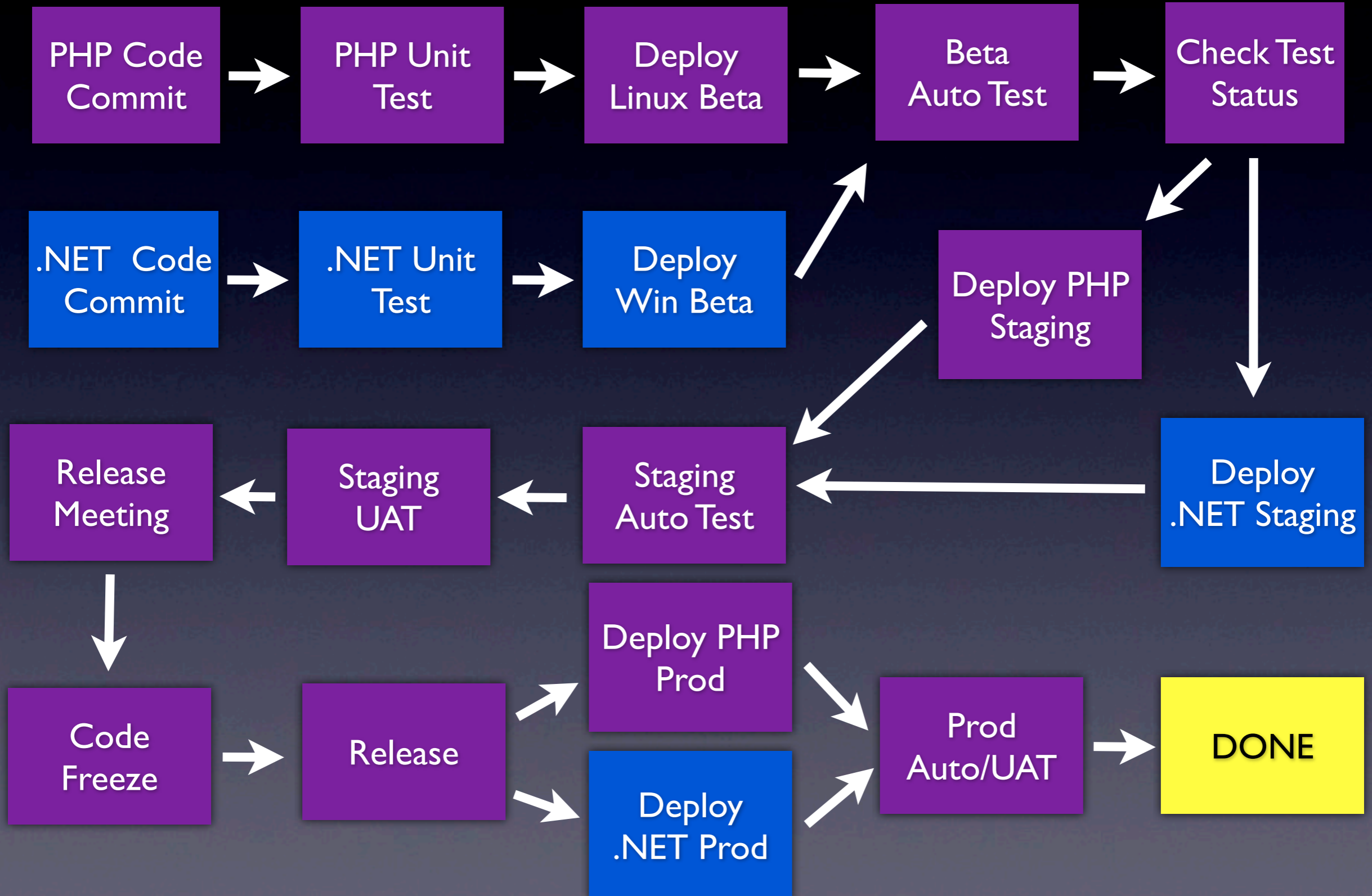
Machine Learning

Data Mart

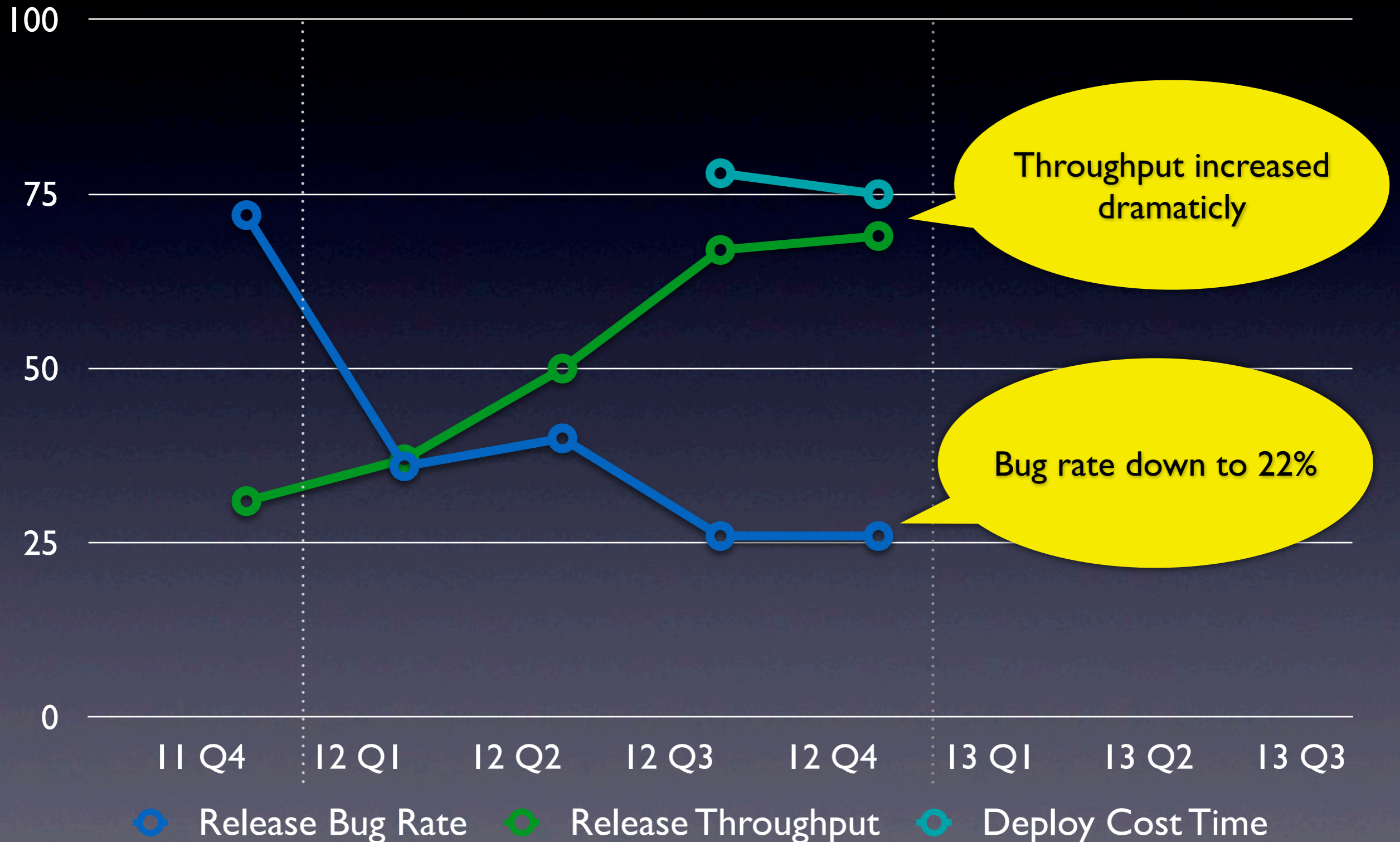
Business Intelligence

CRM

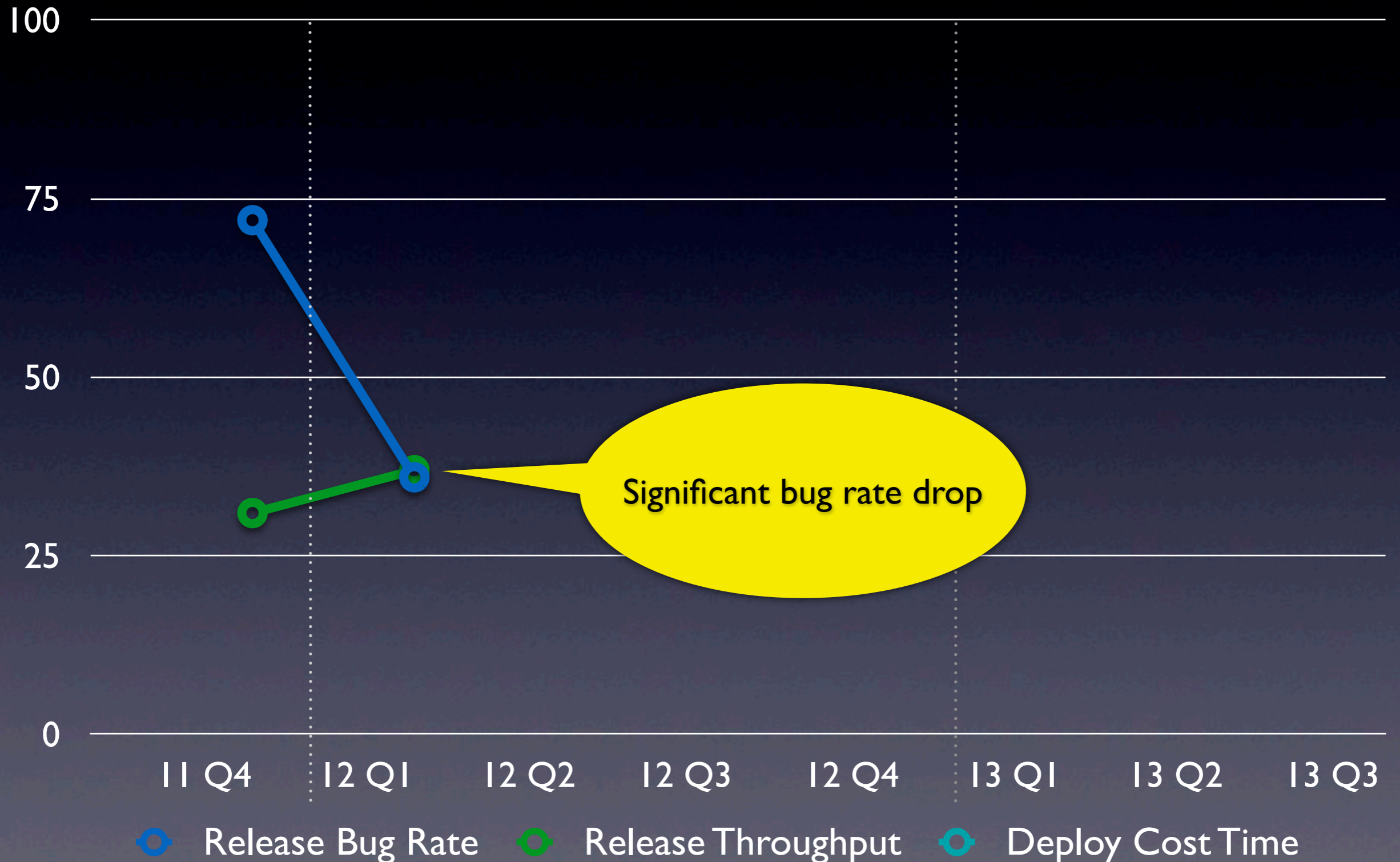
Build and release process



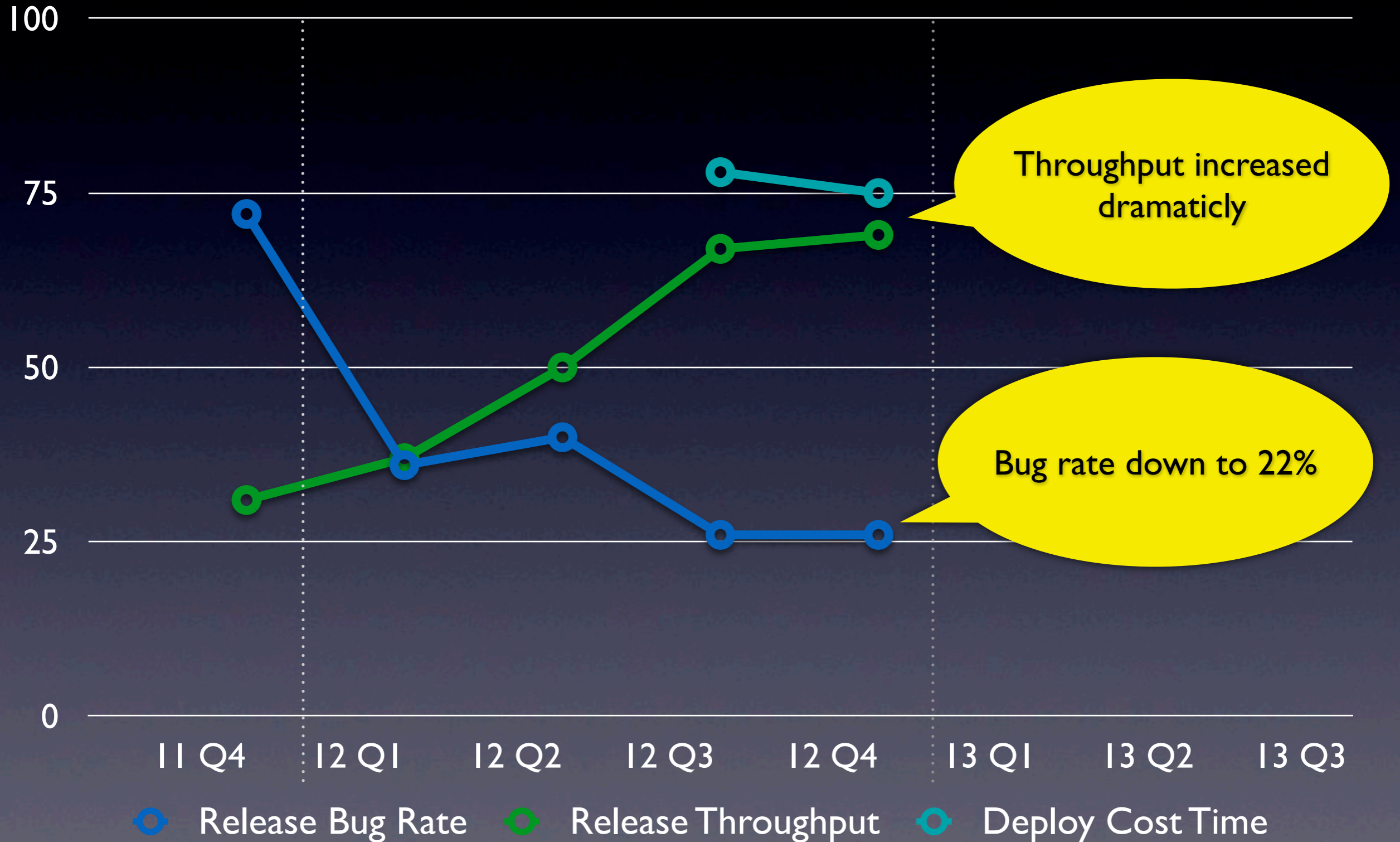
Build CI infrastructure



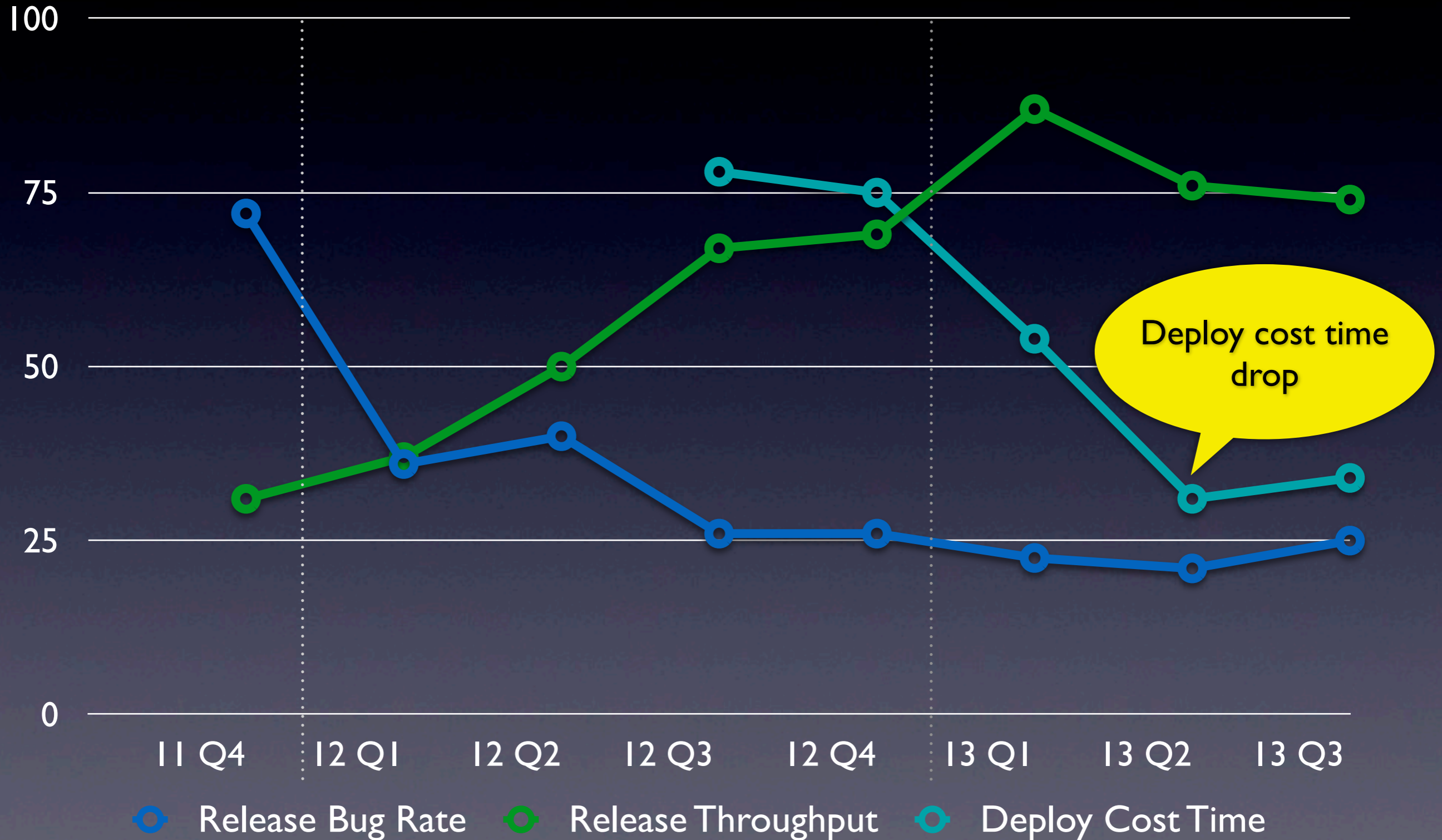
I2Q1 : Create B&R Process



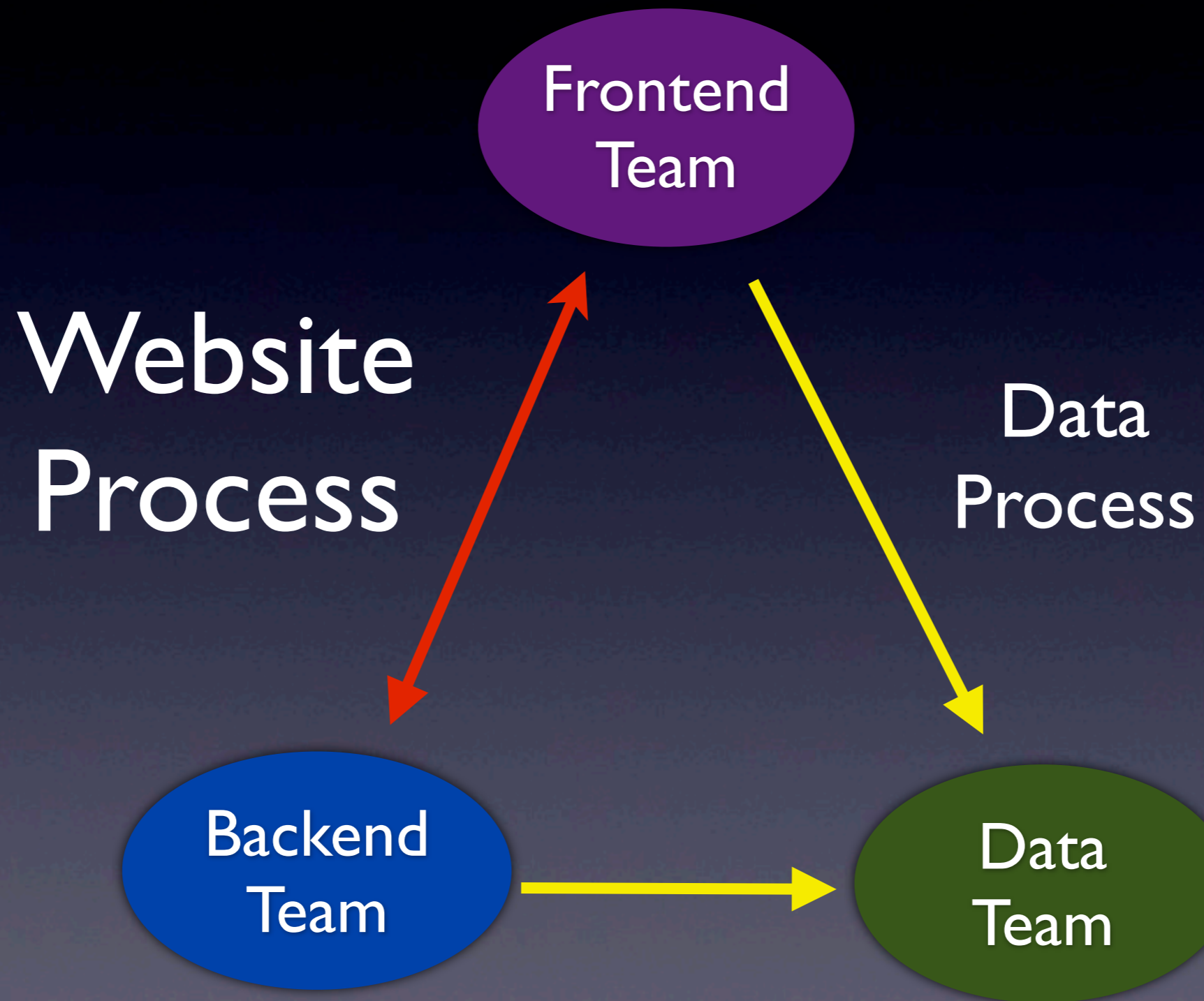
Build CI infrastructure



Automate Everything



Two B&R Processes



Automate Everything

