Introduction
Introduction

Diego: Cloud Foundry’s container runtime
Introduction

Diego: Cloud Foundry’s container runtime

Jan 2014: Started development
Introduction

Diego: Cloud Foundry’s container runtime

Jan 2014: Started development
Sep 2015: Generally available
Introduction

Diego: Cloud Foundry’s container runtime

Jan 2014: Started development
Sep 2015: Generally available
Nov 2016: v1.0.0 after 250K scale
Introduction

Diego: Cloud Foundry’s container runtime

Jan 2014: Started development
Sep 2015: Generally available
Nov 2016: v1.0.0 after 250K scale
May 2017: Retired the DEAs
Topics
Topics

• Diego Components in Cloud Foundry
Topics

• Diego Components in Cloud Foundry
• Core Components and Evolution
Topics

• Diego Components in Cloud Foundry
• Core Components and Evolution
• Updates: Isolation Segments, Instance Credentials
Topics

• Diego Components in Cloud Foundry
• Core Components and Evolution
• Updates: Isolation Segments, Instance Credentials
• Operator Tooling: cfdot
Topics

• Diego Components in Cloud Foundry
• Core Components and Evolution
• Updates: Isolation Segments, Instance Credentials
• Operator Tooling: cfdot
• Next Steps for Diego
Diego in Cloud Foundry
Diego in Cloud Foundry

Cloud Controller
Diego in Cloud Foundry

Cloud Controller

Gorouter
Diego in Cloud Foundry

cf push

Cloud Controller

Gorouter
Diego in Cloud Foundry

cf push

Cloud Controller

Gorouter
Diego in Cloud Foundry

cf push

Cloud Controller

Gorouter
Diego in Cloud Foundry

- Cloud Controller
- Gorouter

[Logos: Apache, Docker, Windows]
Diego in Cloud Foundry

Cloud Controller

Gorouter
Diego in Cloud Foundry

- Cloud Controller
- Gorouter

container execution

- garden
  - docker
  - Windows
Diego in Cloud Foundry

Cloud Controller

Gorouter

container placement & lifecycle

Diego Core

bbs  rep

auctioneer

container execution
garden

SQL DB

consistent datastore
Diego in Cloud Foundry

Cloud Controller

Gorouter

container placement & lifecycle

container execution

garden

Diego Core

bbs  rep

auctioneer

consul

component discovery & coordination

SQL DB

consistent datastore
Diego in Cloud Foundry

- **Cloud Controller**
- **Gorouter**
- **Diego Core**
  - bbs
  - rep
  - auctioneer
- **garden**
- **container execution**
- **container placement & lifecycle**
- **component discovery & coordination**
  - consul
  - locket
- **SQL DB**
  - consistent datastore

- **consistent datastore**
Diego in Cloud Foundry

- **Cloud Controller**
- **Gorouter**
- **SSH**
- **Diego Core**
  - `bbs`
  - `rep`
  - `auctioneer`
- **container placement & lifecycle**
- **garden**
- **container execution**
- **consul**
- **locket**
- **component discovery & coordination**
- **SQL DB**
- **consistent datastore**
Diego Core Components
Diego Core Components
Diego Core Components

Cell

garden
Diego Core Components

- Creates containers
- Runs processes
Diego Core Components

- creates containers
- runs processes
Diego Core Components

- registers cell presence
- controls local garden
- accounts for memory, disk
- caches assets
- creates containers
- runs processes
Diego Core Components

- registers cell presence
- controls local garden
- accounts for memory, disk
- caches assets

API
- bbs

Cell
- rep
- garden

‣ creates containers
‣ runs processes
Diego Core Components

- public API for clients
- enforces lifecycle policy for LRPs and Tasks

- registers cell presence
- controls local garden
- accounts for memory, disk
- caches assets

API
  bbs

Cell
  rep  garden

- creates containers
- runs processes

‣ public API for clients
‣ enforces lifecycle policy for LRPs and Tasks
‣ creates containers
‣ runs processes
Diego Core Components

- public API for clients
- enforces lifecycle policy for LRP Task dependencies
- registers cell presence
- controls local garden
- accounts for memory, disk
- creates containers
- runs processes
- caches assets

API

bbs

Brain

auctioneer

Cell

rep
garden
Diego Core Components

- public API for clients
- enforces lifecycle policy for LRPs and Tasks
- communicates with Cells
- decides optimal placement
- creates containers
- runs processes

API
- bbs

Brain
- auctioneer

Cell
- rep
- garden
- registers cell presence
- controls local garden
- accounts for memory, disk
- caches assets
Diego Core Components

- public API for clients
- enforces lifecycle policy for LRPs and Tasks

API

- bbs
- locket

Brain

- auctioneer

Cell

- rep
- garden

- communicates with Cells
- decides optimal placement
- creates containers
- runs processes

- registers cell presence
- controls local garden
- accounts for memory, disk
- caches assets
Diego Core Components

- public API for clients
- enforces lifecycle policy for LRP's and Tasks

API
- bbs
- locket

Brain
- auctioneer

Cell
- rep
- garden

‣ creates containers
‣ runs processes
‣ communicates with Cells
‣ decides optimal placement
‣ API for locks and presences
‣ expires entries after TTL

‣ registers cell presence
‣ controls local garden
‣ accounts for memory, disk
‣ caches assets
Diego Core Components

- public API for clients
- enforces lifecycle policy for LRP's and Tasks
- compares desired & actual
- fixes discrepancies
- registers cell presence
- controls local garden
- accounts for memory, disk
- caches assets
- creates containers
- runs processes
- API for locks and presences
- expires entries after TTL
- communicates with Cells
- decides optimal placement
- runs processes
# Diego Cell Evolution

<table>
<thead>
<tr>
<th>Cell</th>
</tr>
</thead>
</table>

---
Diego Cell Evolution

Cell

rep

garden-linux

OCI bundle
Diego Cell Evolution

<table>
<thead>
<tr>
<th>Cell</th>
</tr>
</thead>
</table>

rep

garden-linux

OCI bundle

runc
Diego Cell Evolution

Cell

rep

guardian

OCI bundle

runc
Diego Cell Evolution

Cell

rep

guardian

\{ \text{garden-runc} \}

OCI bundle

runc
Diego Cell Evolution

Cell

rep

guardian

garden-runc

OCI bundle

runc

CSI

volume plugins
Diego Cell Evolution

- Cell
- rep
- guardian
- runc
- CSI
  - volume plugins
- CNI
  - network plugin
- OCI bundle
Diego Cell Evolution

<table>
<thead>
<tr>
<th>Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td>rep</td>
</tr>
<tr>
<td>CSI</td>
</tr>
<tr>
<td>volume plugins</td>
</tr>
<tr>
<td>OCI bundle</td>
</tr>
<tr>
<td>guardian</td>
</tr>
<tr>
<td>CNI</td>
</tr>
<tr>
<td>network plugin</td>
</tr>
<tr>
<td>OCI image</td>
</tr>
<tr>
<td>rootfs</td>
</tr>
<tr>
<td>runc</td>
</tr>
</tbody>
</table>
Diego Cell Evolution

Cell

- rep
- CSI
  - volume plugins
- OCI image
- guardian
- CNI
  - network plugin
- grootfs
- OCI bundle
- winc

OCI image

Cloud Foundry Summit 2017
Diego Cell Evolution

**Cell**

- **rep**
- **CSI** volume plugins
- **guardian**
- **OCI bundle** network plugin
- **winc**

**OCI image**

- **grootfs**

*Experimental!*
Placement for Isolation Segments
Placement for Isolation Segments

Cell

Cloud Foundry Summit 2017
Placement for Isolation Segments
## Placement for Isolation Segments

<table>
<thead>
<tr>
<th>Cell</th>
<th>Cell</th>
<th>Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="linux.png" alt="Linux Logo" /></td>
<td><img src="windows.png" alt="Windows Logo" /></td>
<td><img src="green.png" alt="Green Symbol" /></td>
</tr>
</tbody>
</table>
Placement for Isolation Segments

Cell

Cell

Cell

Cell
Placement for Isolation Segments

Apps

<table>
<thead>
<tr>
<th>Cell</th>
<th>Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Linux" /></td>
<td><img src="image" alt="Linux" /></td>
</tr>
<tr>
<td><img src="image" alt="Windows" /></td>
<td><img src="image" alt="Windows" /></td>
</tr>
</tbody>
</table>

## Placement for Isolation Segments

### Apps

<table>
<thead>
<tr>
<th></th>
<th>Cell</th>
<th>Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="docker.png" alt="docker" /></td>
<td><img src="docker.png" alt="docker" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Cell</th>
<th>Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="Linux.png" alt="Linux" /></td>
<td><img src="Green.png" alt="Green" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Cell</th>
<th>Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="Windows.png" alt="Windows" /></td>
<td><img src="Purple.png" alt="Purple" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Cell</th>
<th>Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="Linux.png" alt="Linux" /></td>
<td><img src="Green.png" alt="Green" /></td>
</tr>
</tbody>
</table>
Placement for Isolation Segments

Apps

Cell

Cell

Cell

Cell
Placement for Isolation Segments

Apps

- docker
- Windows
- Linux

Cell with Linux

Cell with Windows

Cell with Linux

Cell with Windows
Placement for Isolation Segments

Apps

Cell

Cell

Cell

Cell
Placement for Isolation Segments

Apps

<table>
<thead>
<tr>
<th>Cell</th>
<th>Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td>docker</td>
<td>docker</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cell</th>
<th>Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows</td>
<td>Windows</td>
</tr>
</tbody>
</table>
Placement for Isolation Segments

Apps

- Cell
  - docker
  - docker

- Cell
  - Windows

- Cell
  - Windows
Placement for Isolation Segments

Apps

Cell

Cell

Cell

Cell

- docker
- docker

- Windows

- Windows
Instance-Identity Credentials
Instance-Identity Credentials

<table>
<thead>
<tr>
<th>Cell</th>
<th>rep</th>
</tr>
</thead>
</table>


Instance-Identity Credentials
Instance-Identity Credentials

Cell

rep

ip 10.255.3.7
Instance-Identity Credentials

<table>
<thead>
<tr>
<th>Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td>rep</td>
</tr>
</tbody>
</table>

| ip   | 10.255.3.7 |
| app  | a83-c2f    |
### Instance-Identity Credentials

<table>
<thead>
<tr>
<th>Cell</th>
<th>rep</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip</td>
<td>10.255.3.7</td>
</tr>
<tr>
<td>app</td>
<td>a83-c2f</td>
</tr>
<tr>
<td>ins</td>
<td>3d5-b3e</td>
</tr>
</tbody>
</table>
Instance-Identity Credentials

Cell

rep

ip 10.255.3.7
app a83-c2f
ins 3d5-b3e
# Instance-Identity Credentials

<table>
<thead>
<tr>
<th>Cell</th>
<th>rep</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip</td>
<td>10.255.3.7</td>
</tr>
<tr>
<td>app</td>
<td>a83-c2f</td>
</tr>
<tr>
<td>ins</td>
<td>3d5-b3e</td>
</tr>
</tbody>
</table>
Instance-Identity Credentials

Cell

rep

<table>
<thead>
<tr>
<th>ip</th>
<th>10.255.3.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>app</td>
<td>a83-c2f</td>
</tr>
<tr>
<td>ins</td>
<td>3d5-b3e</td>
</tr>
</tbody>
</table>
## Instance-Identity Credentials

<table>
<thead>
<tr>
<th>Cell</th>
<th>rep</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip</td>
<td>10.255.3.7</td>
</tr>
<tr>
<td>app</td>
<td>a83-c2f</td>
</tr>
<tr>
<td>ins</td>
<td>3d5-b3e</td>
</tr>
</tbody>
</table>
Instance-Identity Credentials

Cell

rep

ip 10.255.3.7
app a83-c2f
ins 3d5-b3e

Subject: OU=app:a83-c2f/CN=3d5-b3e
Instance-Identity Credentials

Subject: OU=app:a83-c2f/CN=3d5-b3e
Subject Alternative Name:
  IP: 10.255.3.7
  DNS: 3d5-b3e
Instance-Identity Credentials

Subject: OU=app:a83-c2f/CN=3d5-b3e

Subject Alternative Name:
IP: 10.255.3.7
DNS: 3d5-b3e

Validity:
Not Before: 2017-10-11 09:40:00
Not After: 2017-10-12 09:40:00
Instance-Identity Credentials

Subject: OU=app:a83-c2f/CN=3d5-b3e
Subject Alternative Name:
  IP: 10.255.3.7
  DNS: 3d5-b3e
Validity:
  Not Before: 2017-10-11 09:40:00
  Not After: 2017-10-12 09:40:00

} 1 day!
Instance-Identity Credentials

Subject: OU=app:a83-c2f/CN=3d5-b3e

Subject Alternative Name:
  IP: 10.255.3.7
  DNS: 3d5-b3e

Validity:
  Not Before: 2017-10-11 09:40:00
  Not After: 2017-10-12 09:40:00

Serial: a2:3f:49:2b

1 day!
Instance-Identity Credentials

<table>
<thead>
<tr>
<th>Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td>rep</td>
</tr>
</tbody>
</table>

- ip 10.255.3.7
- app a83-c2f
- ins 3d5-b3e
Instance-Identity Credentials

Cell

rep

ip  10.255.3.7
app a83-c2f
ins 3d5-b3e
Instance-Identity Credentials

<table>
<thead>
<tr>
<th>Cell</th>
<th>rep</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip</td>
<td>10.255.3.7</td>
</tr>
<tr>
<td>app</td>
<td>a83-c2f</td>
</tr>
<tr>
<td>ins</td>
<td>3d5-b3e</td>
</tr>
</tbody>
</table>
Instance-Identity Credentials

Subject: OU=app:a83-c2f/CN=3d5-b3e

<table>
<thead>
<tr>
<th>Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td>rep</td>
</tr>
</tbody>
</table>

| ip  | 10.255.3.7 |
| app | a83-c2f |
| ins | 3d5-b3e |
Instance-Identity Credentials

Subject: OU=app:a83-c2f/CN=3d5-b3e
Subject Alternative Name:
IP: 10.255.3.7
DNS: 3d5-b3e
Instance-Identity Credentials

Subject: OU=app:a83-c2f/CN=3d5-b3e

Subject Alternative Name:
- IP: 10.255.3.7
- DNS: 3d5-b3e

Validity:
- Not Before: 2017-10-12 08:40:00
- Not After: 2017-10-13 08:40:00

Cell

rep

ip 10.255.3.7
app a83-c2f
ins 3d5-b3e

📜 🏴‍☠️ 🔑
Instance-Identity Credentials

Subject: OU=app:a83-c2f/CN=3d5-b3e

Subject Alternative Name:
IP: 10.255.3.7
DNS: 3d5-b3e

Validity:
Not Before: 2017-10-12 08:40:00
Not After: 2017-10-13 08:40:00

Serial: f7:5d:53:c1
Instance-Identity Credentials

Cell

rep

ip 10.255.3.7
app a83-c2f
ins 3d5-b3e
Secure Direct App Communication
Secure Direct App Communication
Secure Direct App Communication

gorouter

CF
Secure Direct App Communication

gorouter
cf
frontend-new
Secure Direct App Communication

- gorouter
- frontend-new
- frontend-old
- CF
Secure Direct App Communication

gorouter

frontend-new

frontend-old

backend

CF
Secure Direct App Communication

- gorouter
- frontend-new
  - app a83-c2f
- frontend-old
- backend

CF
Secure Direct App Communication

- frontend-new
  - app a83-c2f
- frontend-old
  - app c63-2de
- gorouter
- CF
- backend

[Diagram showing the flow of applications and services with connections between gorouter, CF, frontend-new, frontend-old, and backend.]
Secure Direct App Communication

- **gorouter**
  - frontend-new
    - app a83-c2f
  - frontend-old
    - app c63-2de

- **backend**
  - ✅:a83-c2f
Secure Direct App Communication

neato.com ➔ gorouter

CF

frontend-new
app a83-c2f

frontend-old
app c63-2de

backend
✅:a83-c2f
Secure Direct App Communication

neato.com → gorouter

- frontend-new
  - app a83-c2f

- frontend-old
  - app c63-2de

- backend
  - 👍:a83-c2f
Secure Direct App Communication

neato.com ➔ gorouter ➔ frontend-new
app a83-c2f ➔ backend
✅:a83-c2f

CF

frontend-old
app c63-2de
Secure Direct App Communication

neato.com

gorouter

frontend-new
app a83-c2f

frontend-old
app c63-2de

backend

✅:a83-c2f

CF
Secure Direct App Communication

neato.com

gorouter

frontend-new
app a83-c2f

frontend-old
app c63-2de

backend
✅:a83-c2f

CF
Secure Direct App Communication

- **neato.com**
  - gorouter
  - frontend-new
    - app a83-c2f
  - frontend-old
    - app c63-2de
  - backend
    - :a83-c2f
Secure Direct App Communication

- gorouter
- frontend-new
  - app a83-c2f
- frontend-old
  - app c63-2de
- backend
  - :a83-c2f
Secure Direct App Communication

old.neato.com → gorouter → frontend-new
  app a83-c2f
frontend-old
  app c63-2de
 backend
  ✅:a83-c2f

CF
Secure Direct App Communication

old.neato.com

gorouter

frontend-new
app a83-c2f

frontend-old
app c63-2de

backend
✅:a83-c2f

CF
Secure Direct App Communication

old.neato.com

gorouter

frontend-new
app a83-c2f

frontend-old
app c63-2de

backend
✓:a83-c2f
Secure Direct App Communication

old.neato.com → gorouter → frontend-new
  app a83-c2f → backend
  ✓:a83-c2f

CF → frontend-old
  app c63-2de
Secure Direct App Communication
Secure Direct App Communication

- old.neato.com
  - gorouter
  - frontend-new
    - app a83-c2f
  - frontend-old
    - app c63-2de
  - backend
    - :a83-c2f
Secure Direct App Communication

github.com/emalm/tls-example-apps
cfdot: Diego Operator Toolkit
cfdot: Diego Operator Toolkit

- cfdot: command-line tool for Diego
cfdot: Diego Operator Toolkit

• cfdot: command-line tool for Diego
  • available now: BBS (LRPs, Tasks) and Locket APIs
cfdot: Diego Operator Toolkit

• cfdot: command-line tool for Diego
  • available now: BBS (LRPs, Tasks) and Locket APIs
  • up next: Cell API, environment overview
cfdot: Diego Operator Toolkit

- cfdot: command-line tool for Diego
  - available now: BBS (LRPs, Tasks) and Locket APIs
  - up next: Cell API, environment overview
- Motivation: these APIs not `curl`-friendly (protobuf RPC)
**cfdot: Diego Operator Toolkit**

- **cfdot**: command-line tool for Diego
  - available now: BBS (LRPs, Tasks) and Locket APIs
  - up next: Cell API, environment overview
- Motivation: these APIs not `curl`-friendly (protobuf RPC)
- Output: stream of JSON objects on stdout
**cfdot: Diego Operator Toolkit**

- **cfdot**: command-line tool for Diego
  - available now: BBS (LRPs, Tasks) and Locket APIs
  - up next: Cell API, environment overview
- Motivation: these APIs not *curl*-friendly (protobuf RPC)
- Output: stream of JSON objects on stdout
- BOSH job: deploys **cfdot** and **jq** binaries, adds to **PATH**
cfdot: Examples
cfdot: Examples

Example: Count of app instance states
Example: Count of app instance states

```bash
$ cfdot actual-lrp-groups | jq .instance \
  | jq -s -r 'group_by(.state)[] | "\([0].state): \(length)"
```
**cfdot: Examples**

**Example: Count of app instance states**

```bash
$ cfdot actual-lrp-groups | jq .instance \
  | jq -s -r 'group_by(.state)[] | "\([0].state): \(length)"
```

- **CRASHED:** 38
- **RUNNING:** 288
- **UNCLAIMED:** 3


**cfdot: Examples**

**Example: Count of app instance states**

```
$ cfdot actual-lrp-groups | jq .instance \
  | jq -s -r 'group_by(.state)[] | "\([0].state): \(length)"'
```

<table>
<thead>
<tr>
<th>State</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRASHED</td>
<td>38</td>
</tr>
<tr>
<td>RUNNING</td>
<td>288</td>
</tr>
<tr>
<td>UNCLAIMED</td>
<td>3</td>
</tr>
</tbody>
</table>

**Example: Find app guid and index at IP and port**
**cfdot: Examples**

**Example: Count of app instance states**

```bash
$ cfdot actual-lrp-groups | jq .instance  \
    | jq -s -r 'group_by(.state)[] | "\([0].state): \(length)"'
```

CRASHED: 38
RUNNING: 288
UNCLAIMED: 3

**Example: Find app guid and index at IP and port**

```bash
$ cfdot actual-lrp-groups | jq -r '.instance'
  ' | select(.address=="10.10.50.34" and (.ports[].host_port'
  ' | contains(61077))) | "\(.process_guid[0:36]) \(.index)"
```
**cfdot: Examples**

**Example: Count of app instance states**

```
$ cfdot actual-lrp-groups | jq .instance \
   | jq -s -r 'group_by(.state)[] | "\([0].state): \(length)"
```

<table>
<thead>
<tr>
<th>CRASHED: 38</th>
</tr>
</thead>
<tbody>
<tr>
<td>RUNNING: 288</td>
</tr>
<tr>
<td>UNCLAIMED: 3</td>
</tr>
</tbody>
</table>

**Example: Find app guid and index at IP and port**

```
$ cfdot actual-lrp-groups | jq -r '.instance'\n' | select(.address=="10.10.50.34" and (.ports[].host_port'\n' | contains(61077)) | "\(.process_guid[0:36]) \(.index)"
```

```
ff6d424f-9edc-478c-b591-cdd3a1498460 0
```
Focus for Next Steps
Focus for Next Steps

• Supporting zero-downtime/rolling app updates
Focus for Next Steps

- Supporting zero-downtime/rolling app updates
- Better placement of large apps with less headroom
Focus for Next Steps

• Supporting zero-downtime/rolling app updates
• Better placement of large apps with less headroom
• Automatic back-off of stressed cells, cluster
Focus for Next Steps

• Supporting zero-downtime/rolling app updates
• Better placement of large apps with less headroom
• Automatic back-off of stressed cells, cluster
• Diego v2: require component TLS, require Locket?
Focus for Next Steps

• Supporting zero-downtime/rolling app updates
• Better placement of large apps with less headroom
• Automatic back-off of stressed cells, cluster
• Diego v2: require component TLS, require Locket?
• Continued stability and security improvements
Questions?
Thank you!

emalm@pivotal.io · github.com/emalm · @emalm, #diego in CF OSS Slack

github.com/cloudfoundry/diego-release