

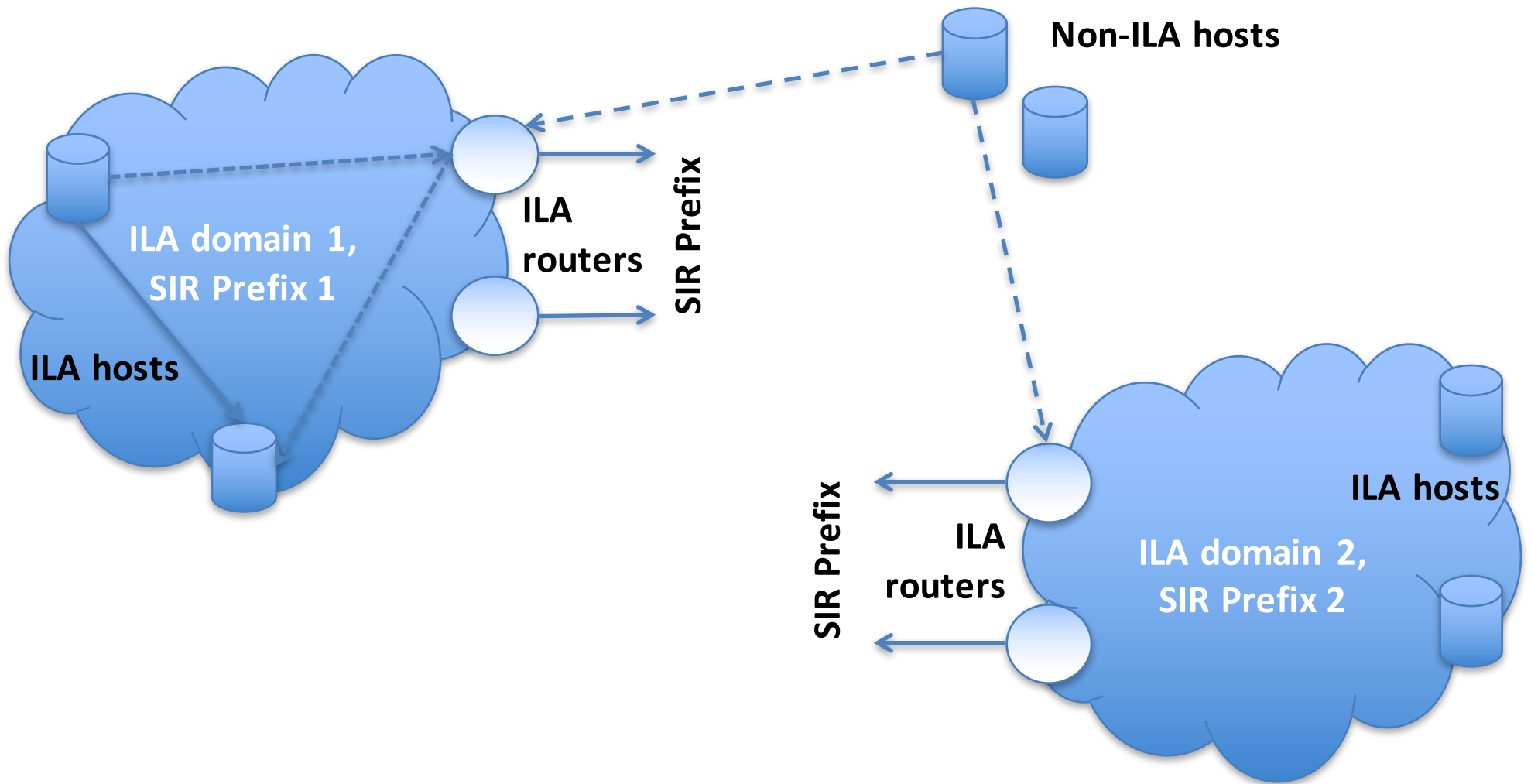
BGP ILA AFI

Petr Lapukhov,
Facebook,
petr@fb.com

Crash course of ILA

- Identifier-Locator Addressing, based on ILNP
 - Adapted for data-center use-case
- **IPv6 = <64-bit locator>:<64-bit Identifier>**
 - Locators – topologically significant
 - Identifiers – application names
- Locators resolved dynamically
- Identifiers can change locators
- **SIR = Standard Identifier Representation**
 - Locator (/64) visible to non-ILA hosts

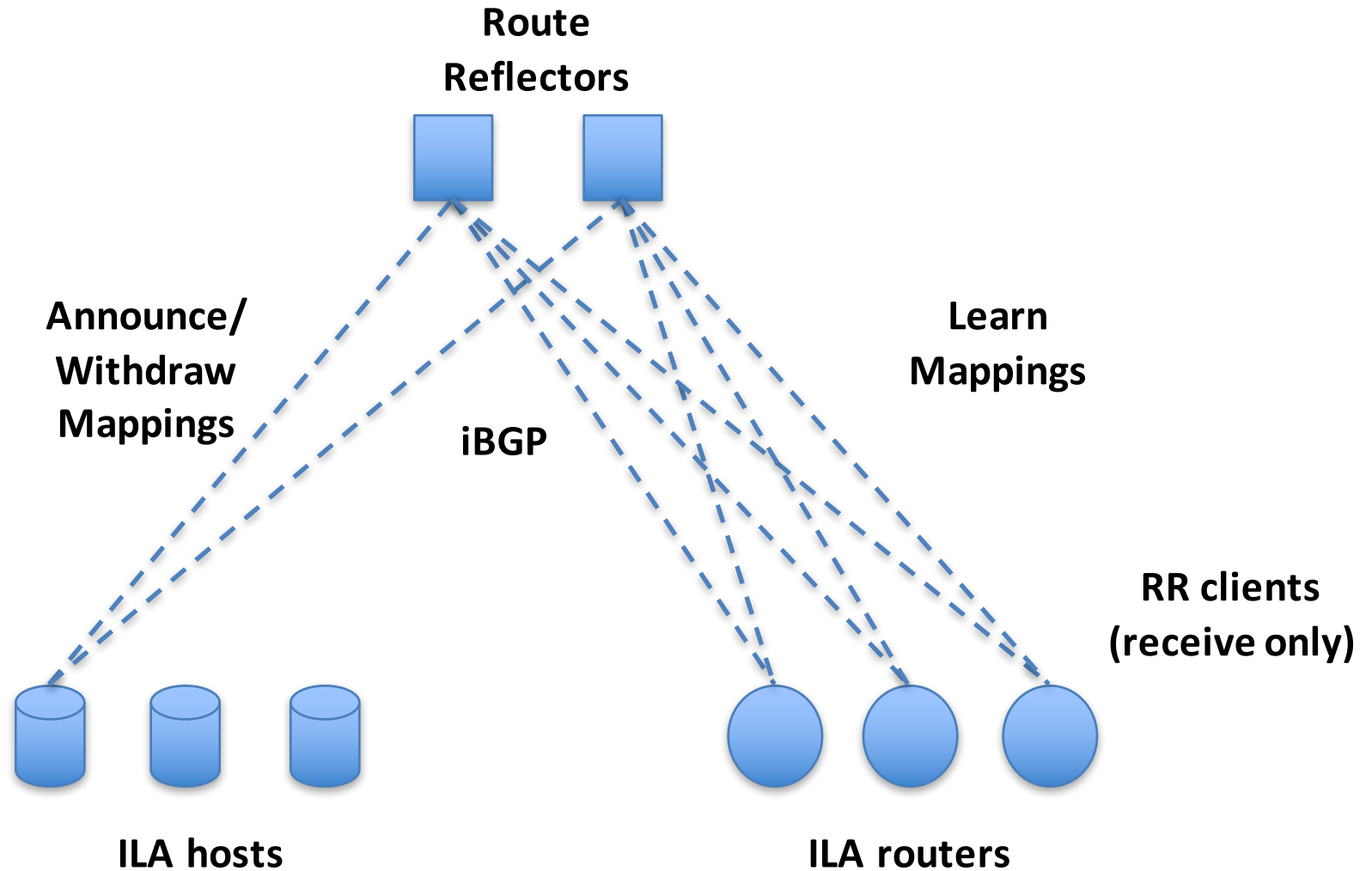
ILA terminology



Crash course of ILA (cont.)

- Every host maintains ILA mapping table
 - Built either via “push” or “pull” method
 - **Pull**: on demand, built via ILA router messages
 - **Push**: populated via BGP messages
- ILA routers
 - Maintain mappings for the domain
 - Learn of ILA mappings from ILA hosts
 - BGP could be used to propagate active mappings

BGP push: Distribution topology (2)



Why BGP, again?

- Re-using existing protocol seems attractive
- ILA routers could be hardware appliances
- BGP implementations available for hosts
 - *nix, Windows
- BGP is known to scale to few million prefixes
- Easy to extend, simple changes

BGP ILA AFI

- AFI: ILA, SAFI: Unicast, Multicast

```
+-----+
| Address Family Identifier (2 octets) |
+-----+
| Subsequent Address Family Identifier (1 octet) |
+-----+
| Length of Next Hop Address (1 octet, set to "8") |
+-----+
| Locator value (8 octets) |
+-----+
| Reserved (1 octet), must be zero |
+-----+
| Length of NLRI field (2 octets, multiple of 8) |
+-----+
| Identifiers (variable, 8 octets each) |
+-----+
```

MP_REACH_NLRI

Inter-domain mapping

- Identifiers do not cross domains
- Talking H2H might be beneficial
 - Shortcut, offload ILA routers
- Need to exchange cross-domain mappings
- Similar to VPN-IPvX AFI, use VPN-ILA AFI
- Use SIR as the RD

Inter-domain mapping

```
+-----+
| Address Family Identifier (2 octets) |
+-----+
| Subsequent Address Family Identifier (1 octet) |
+-----+
| SIR prefix (8 octets) |
+-----+
| Length of Next Hop Address (1 octet, set to "8") |
+-----+
| Locator value (8 octets) |
+-----+
| Reserved (1 octet), must be zero |
+-----+
| Length of NLRI field (2 octets, multiple of 8) |
+-----+
| Identifiers (variable, 8 octets each) |
+-----+
```

MP_REACH_NLRI

Summary

- Very simple extensions
- Model existing BGP mechanics
- Allows re-using BGP deployments/operations
- Scales to few millions of mappings