Dependability of All-IP Networks

Workshop in Espoo, Finland, May 2006

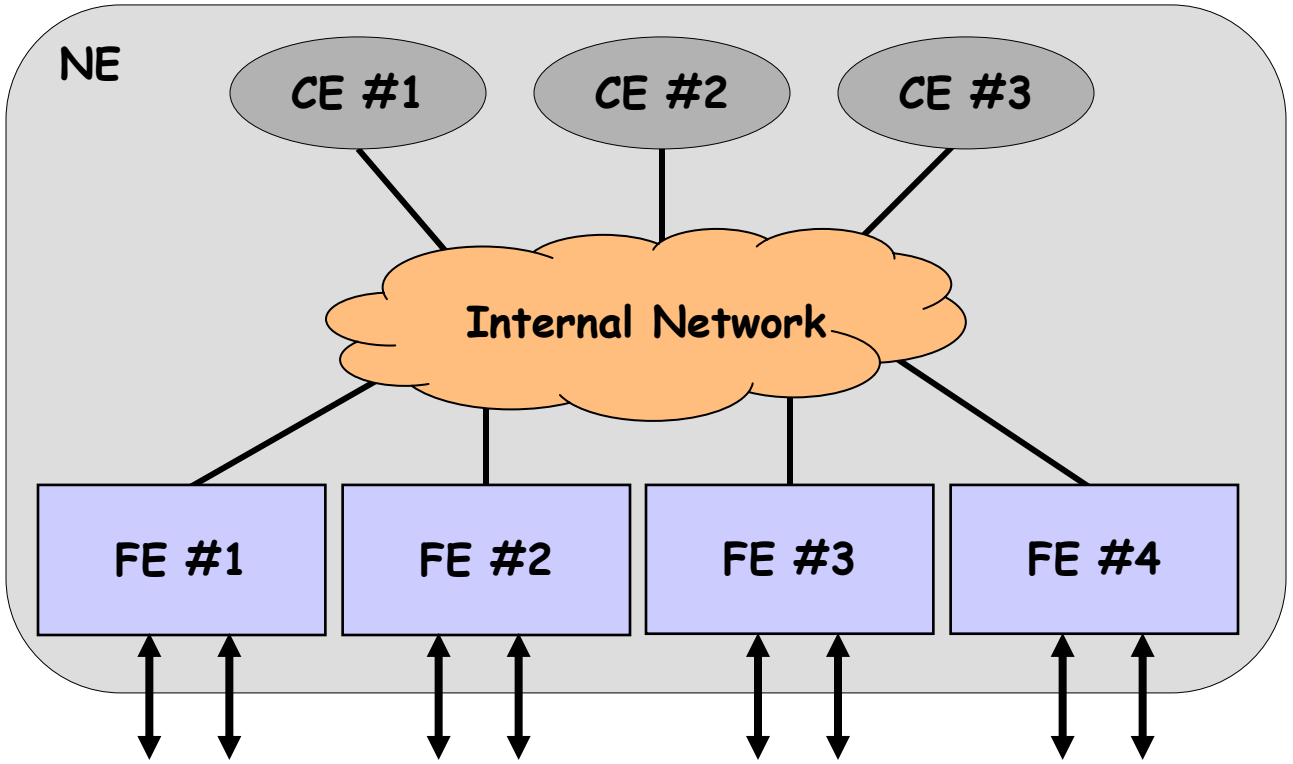
Distributed Architectures for Router Control Plane Robustness

Markus Hidell
Peter Sjödin

Background

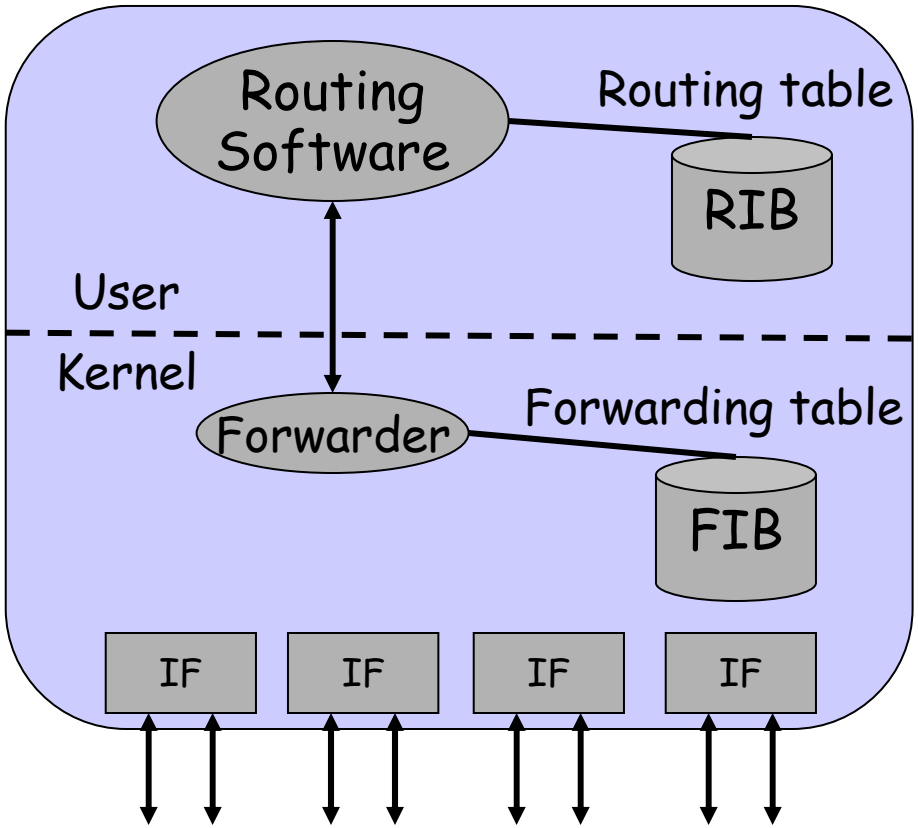
- Research on distributed router architectures
- Increased requirements on IP routers
  - New services (VPN, IPv6, …)
  - New protocols (MPLS, RSVP, …)
  - More features for existing protocols (BGP features)
- Augmenting router functionality affects
  - Control plane robustness
  - Control plane performance
  - Forwarding plane (but that is outside today's scope)
- Can modularization and decentralization in a distributed router architecture improve robustness?

Decentralization and Modularization



- Network system composed of modules mapped onto different processing elements
- Physical separation of forwarding and control elements
- Open well-defined communication interfaces—ForCES

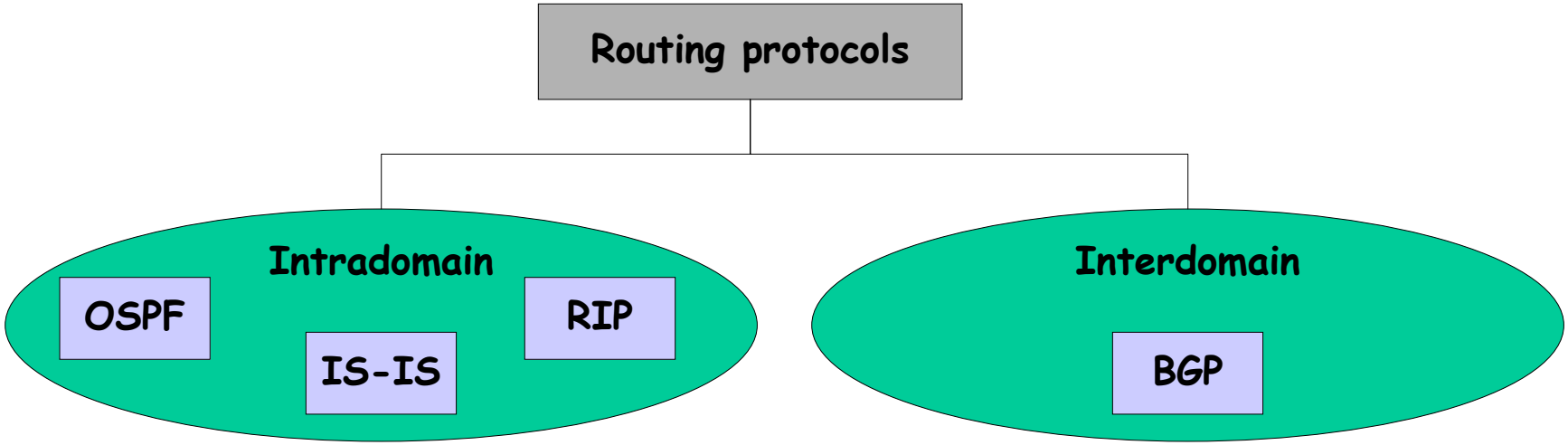
Monolithic Router Control Plane



- Support increasing number of protocols
- Routing protocols among the most important functions

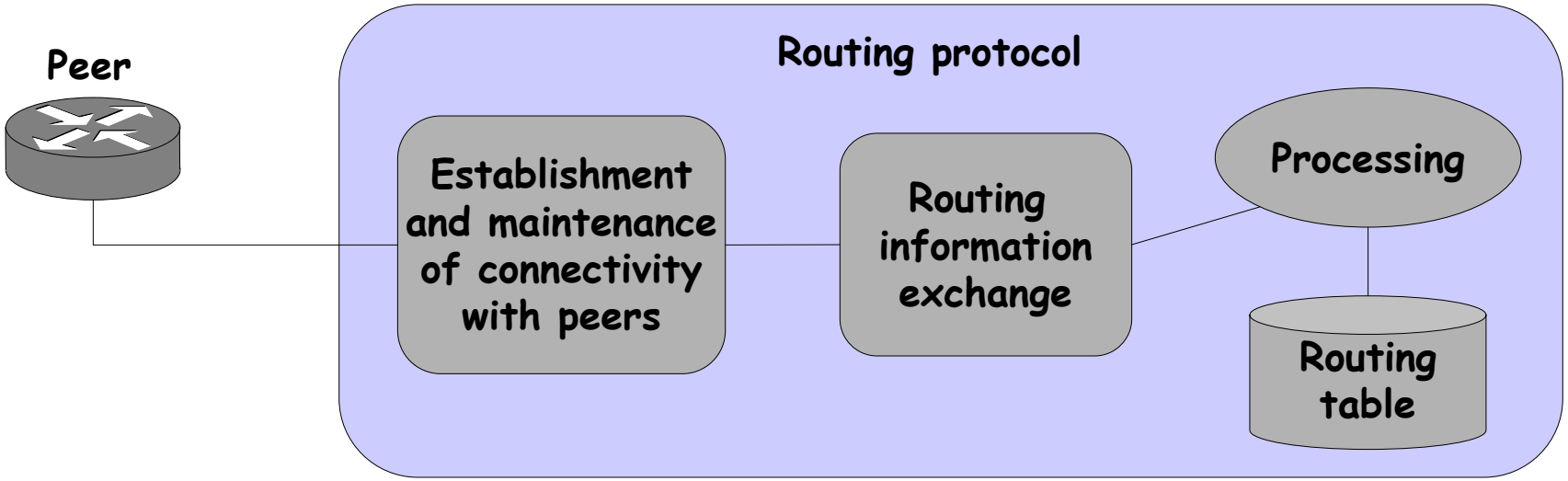
- Control plane traditionally centralized and monolithic
  - Monolithic, one single image, multiple processes share memory space
  - Dual control processors—primary and back-up
- Recent advances (Cisco, Juniper)
  - Modularity, memory protection between processes, threads, multiple route processors
  - Typically: 2 route processors where one is active and one passive

Routing Protocols—Find Best Path



Protocol	What to send	Recipients	Path computation
Distance Vector (RIP)	Routing table	Directly connected neighbor routers	Performed incrementally, no global view
Link state (OSPF, IS-IS)	Link state advertisements	All participating routers	Complete delivery tree using Dijkstra, global view

Decomposing a Routing Protocol

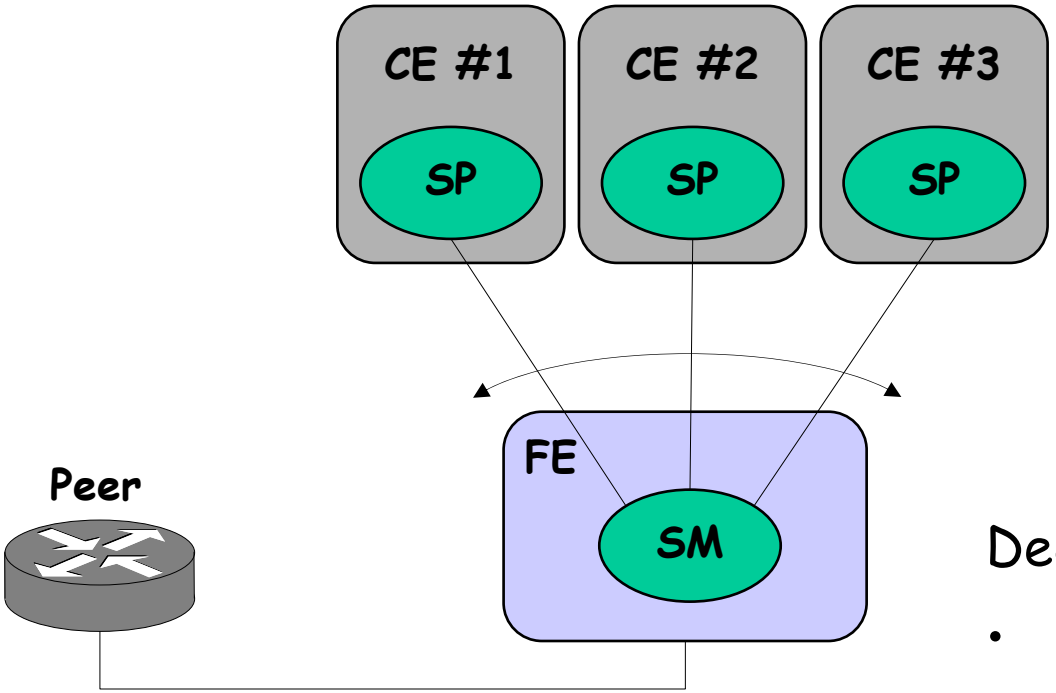


- BGP
  - Establishment of BGP sessions with peers
  - Exchange UPDATE messages
  - Processing UPDATEs and execute *decision process* to select best routes
- OPSF
  - Hello protocol to establish connectivity with peers
  - Exchange and Flooding protocols
  - Processing link state updates and compute Dijkstra's algorithm

Routing Dependency

- Can we exploit modularization and decentralization to improve routing reliability?
- Failures in hardware and software components
  - Redundancy and replication
- Performance aspects
  - Routing computations may potentially lock up routers
    - Sudden changes in topology, route flaps, restarts, DoS attacks, …
  - Reduced vulnerability through distributed computations

Modularization and Decentralization



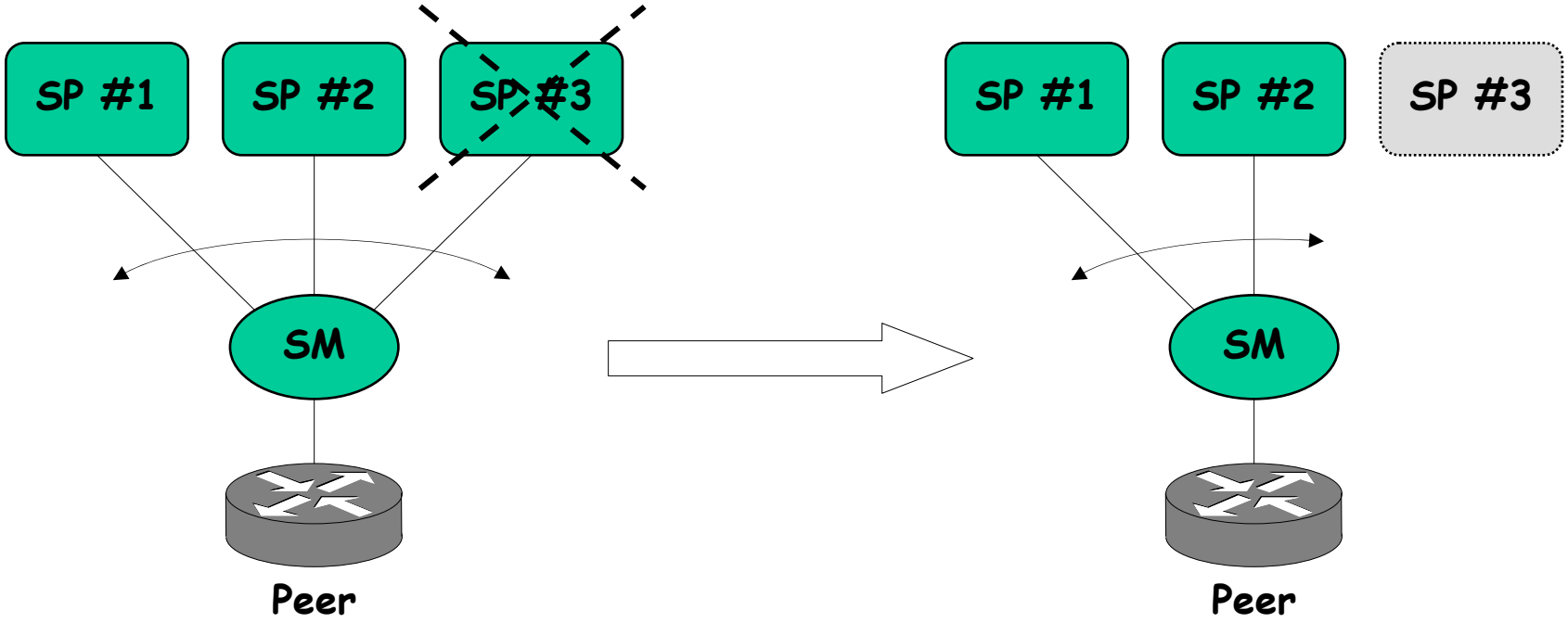
Modularization

- Service process: processes routing information
- Session manager: maintains connectivity with peers

Decentralization

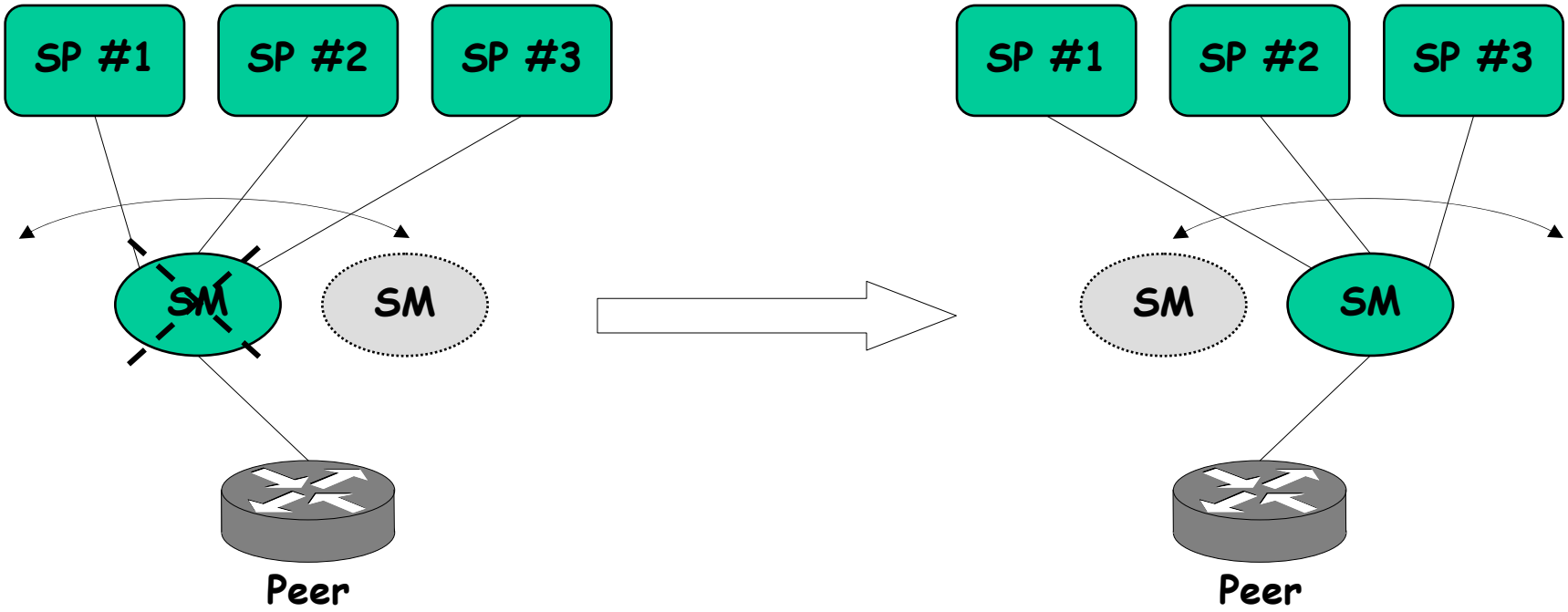
- Map service process modules onto control elements
- Map session managers onto forwarding elements
- Session manager distributes workload over service processes

Service Process Operation



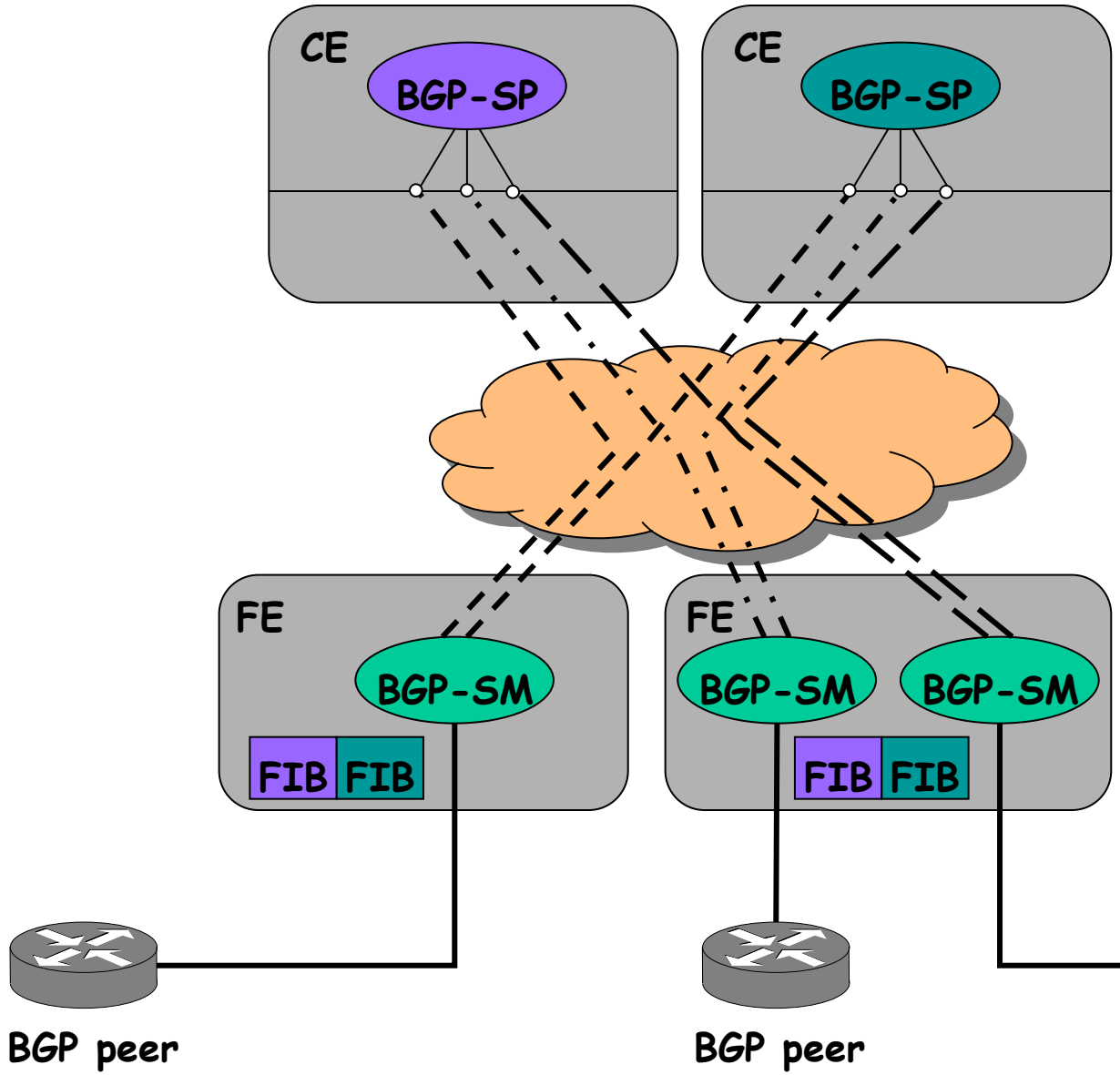
- Collaborating SPs have same capabilities
- Workload can be dynamically redistributed over SPs
- Workload distribution controlled by SM
- External peer should not be affected

Session Manager Operation



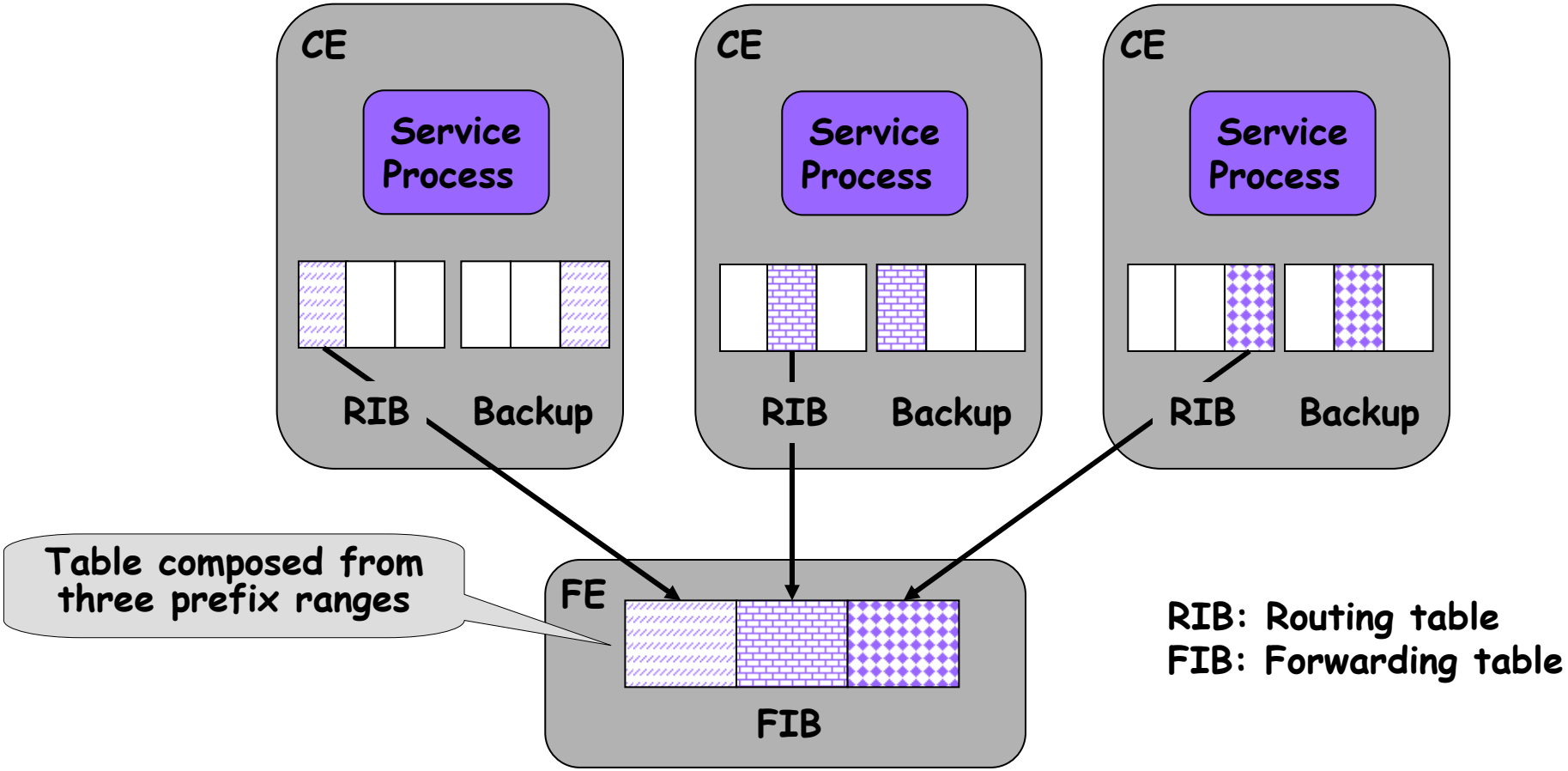
- SM may be replicated for redundancy purposes
- Fail-over without disturbing external peer
  - Depends on transport mechanism and session level state information
  - UDP- or IP-based "sessions" straight-forward (e.g., OSPF)
  - TCP-based sessions more complex (e.g., BGP)
    - Migratory TCP, FT-TCP, etc for seamless migration of TCP connections

A Decentralized Modular BGP



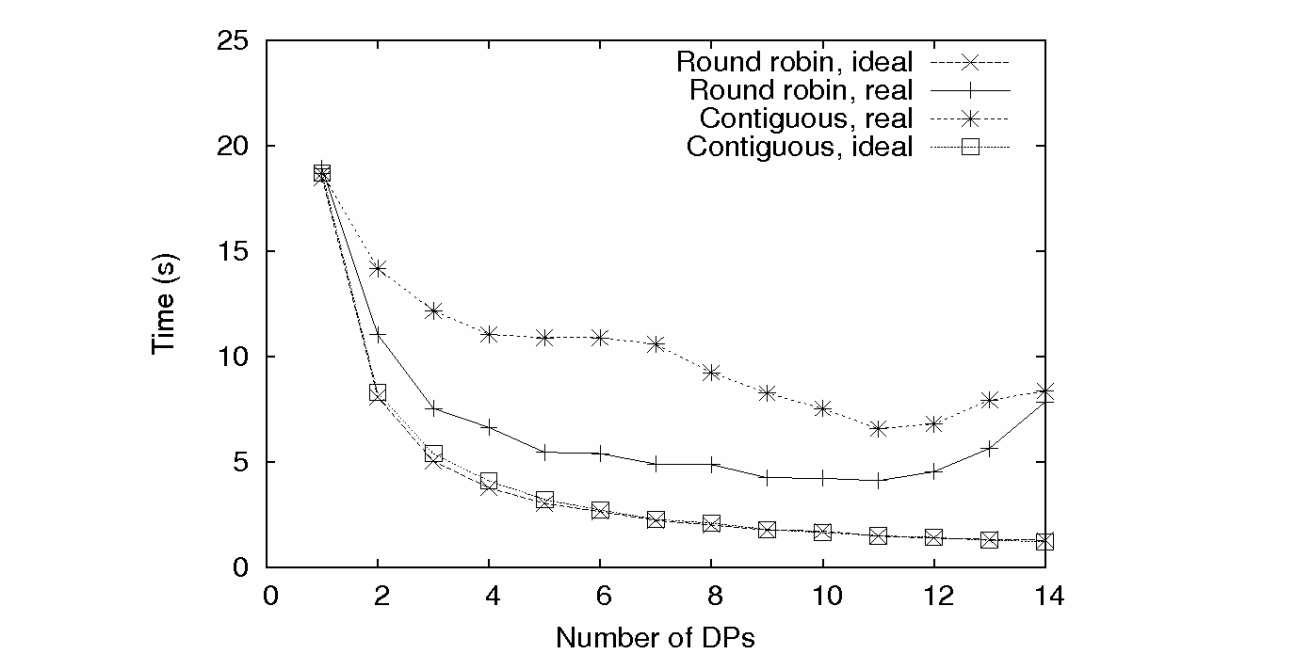
- BGP FSM running on SM
- SM dispatches BGP UPDATEs to SPs
- SPs process BGP UPDATES and compute RIB
- SPs updates FIBs on FEs
- Improved robustness and performance

BGP Load Sharing and Redundancy



- Different SPs responsible for different address prefix ranges
  - Share computational burden
  - Active for one range of prefixes, passive for another → fast failover

Performance Measurements



- Contiguous route assignments to SPs
- Interleaved route assignments to SPs
- BGP input filters used to implement assignment policies
- For real data, and #SPs > 11, the Session manager gets saturated and the total performance drops

Conclusions

- Proposed a distributed control plane for decentralized modular router architectures
- Allows for redundancy and functional replication
  - Further work: failure management support mechanisms in internal protocols
- Applied this approach on BGP
  - Performance improvements
  - Reduce computational load can improve reliability

Thank You for Listening!