

# Ordered FIB Updates

---

draft-francois-ordered-fib-01.txt

*Pierre Francois*  
Olivier Bonaventure  
Mike Shand  
Stefano Previdi  
Stewart Bryant

# Outline

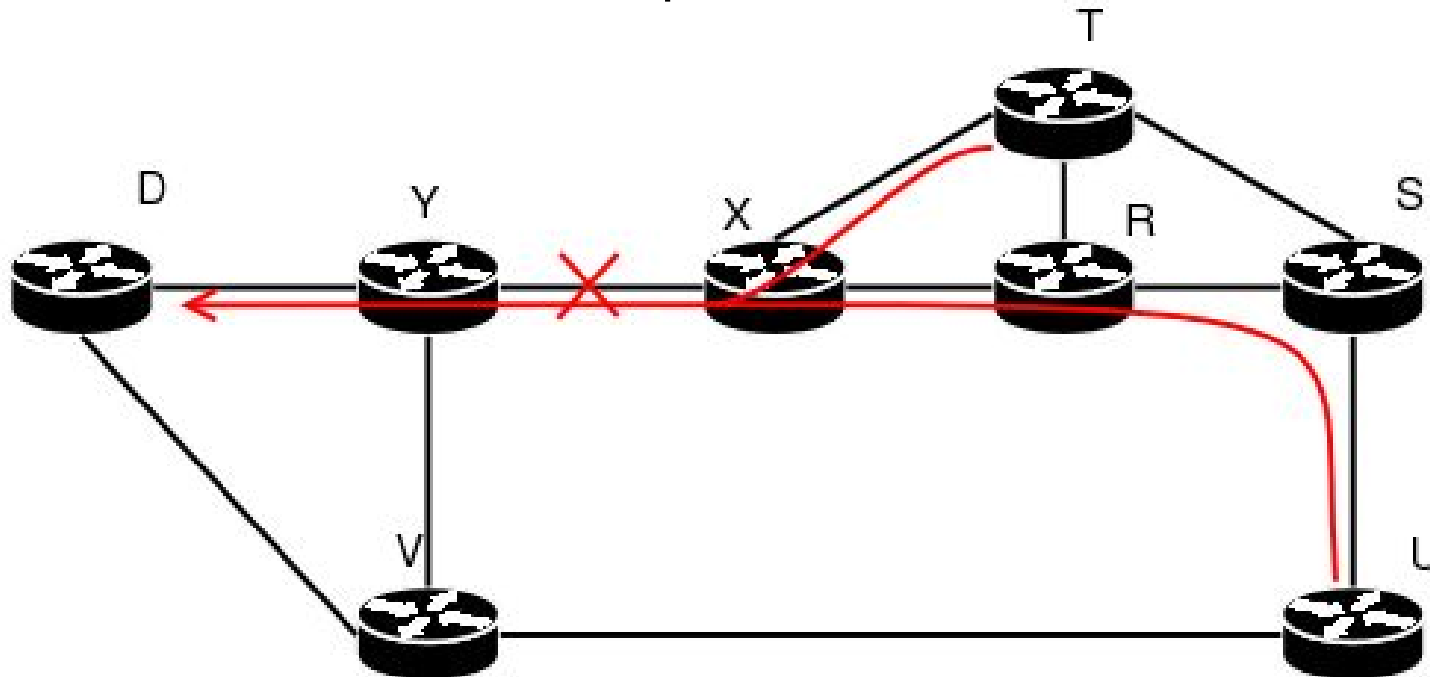
---

- Quick reminder on ordered FIB updates
- Convergence time of ordered FIB updates
- Comparison with PLSN
- Conclusion

# Ordered FIB Updates

## Principle

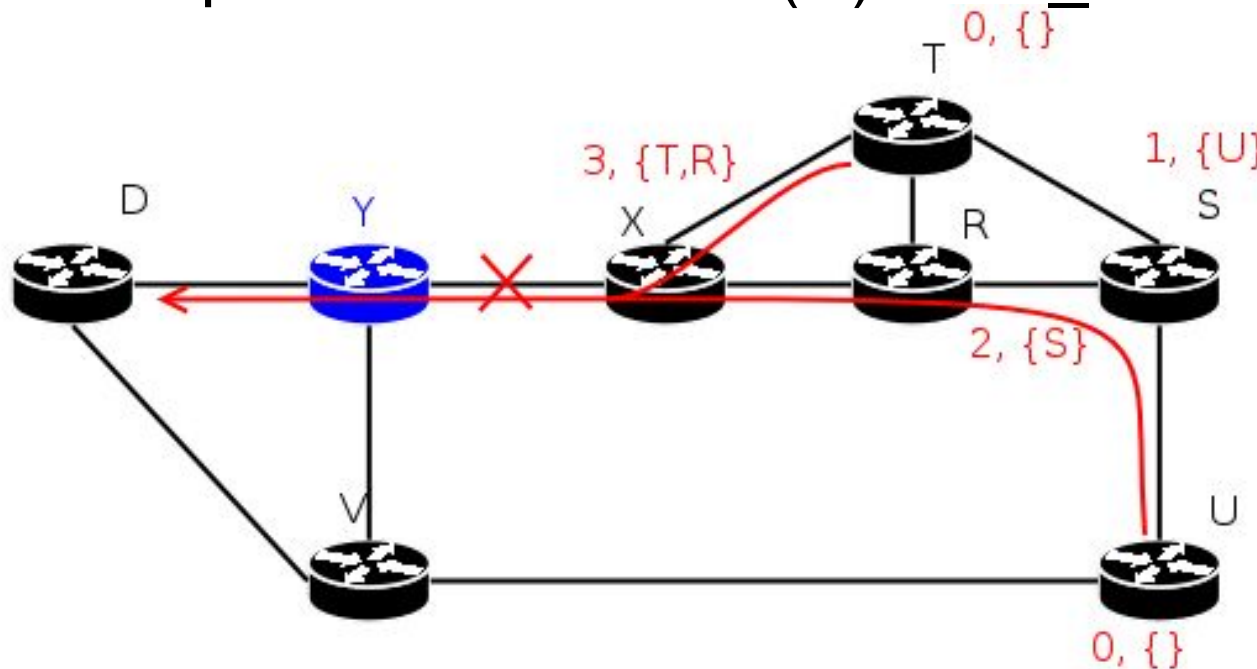
- Order FIB updates on the routers to avoid transient loops
- Considering the removal of a link  $X \rightarrow Y$ 
  - A router **R updates its FIB after** the routers that use itself to reach the removed link ( $X \rightarrow Y$ ).
    - R updates its FIB after S and U
    - T does not wait to update its FIB
    - X is the last router to update its FIB



# Ordered FIB Updates

How to determine the correct ordering ?

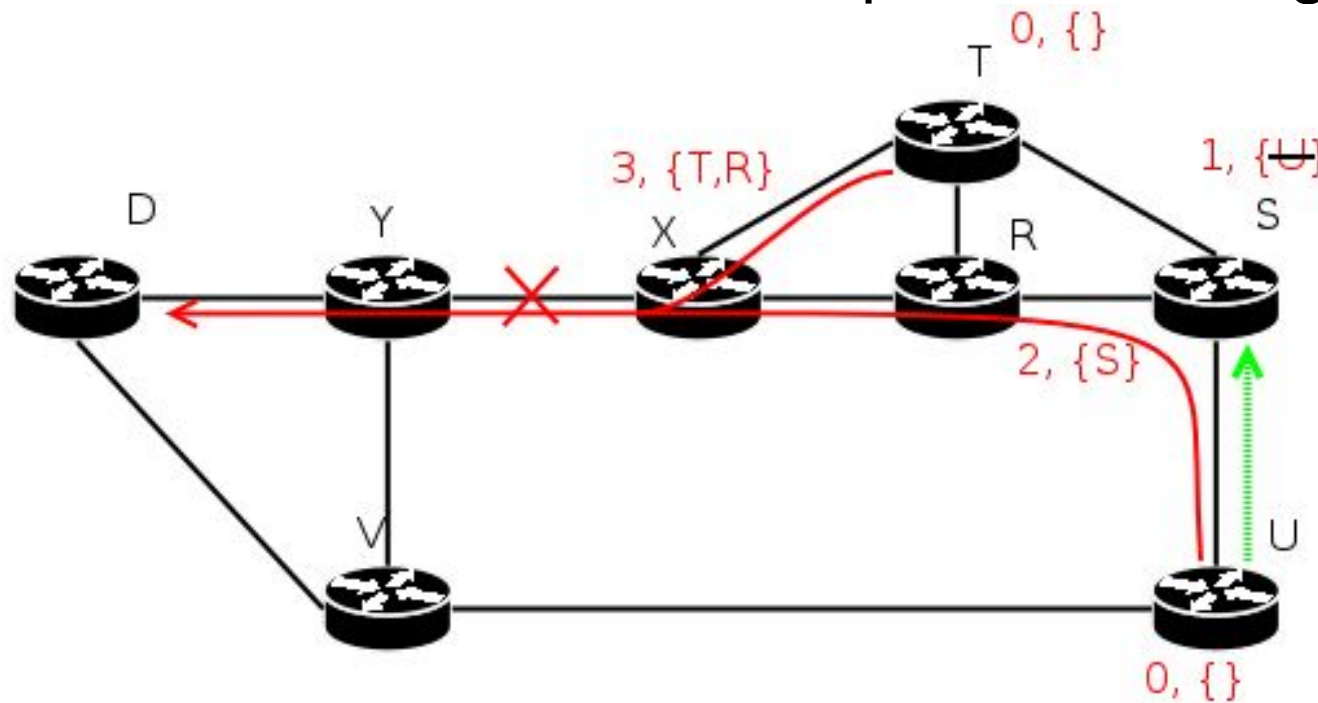
- Each router computes a rank
  - rSPF rooted at **Y** gives the shortest paths to **Y**
  - During rSPF computation
    - R finds its Rank,  $\text{depth}(R, \text{rSPT}(X \rightarrow Y))$ 
      - max (hop) length among paths to R used to reach Y
    - R finds the **set of neighbors** that use it to reach  $(X \rightarrow Y)$ 
      - **Waiting List** of R (used to shortcut the rank)
  - R's FIB update time is  $\text{Rank}(R) * \text{MAX\_FIB}$



# Ordered FIB Updates

## Completion messages

- When a router updates its FIB
  - It sends a **completion message** to its old nexthops for X
- When a router receives a completion message
  - It removes the sender from its Waiting List
- When R's Waiting List becomes empty
  - R can update its FIB and send its completion message
- Rank timer recovers from lost completion messages



# Simulation results

## Ordered FIB convergence time

---

- The time to perform an ordered FIB update after a link-state change, by considering
  - The **flooding** of the link-state packet across the network
    - Link delay
    - LSP processing in the router (4 msec)
  - The computation time of the (r)SPT (**once LSP is received**)
    - This is assumed to take 200 msec
  - The time required to **update a FIB**
    - 100  $\mu$ sec / prefix (measured on Cisco 12k)
  - The time required to process a completion message
    - This is assumed to take 4 msec
- Convergence is reached when all FIBs have been updated and no more completion messages are sent

# First case study

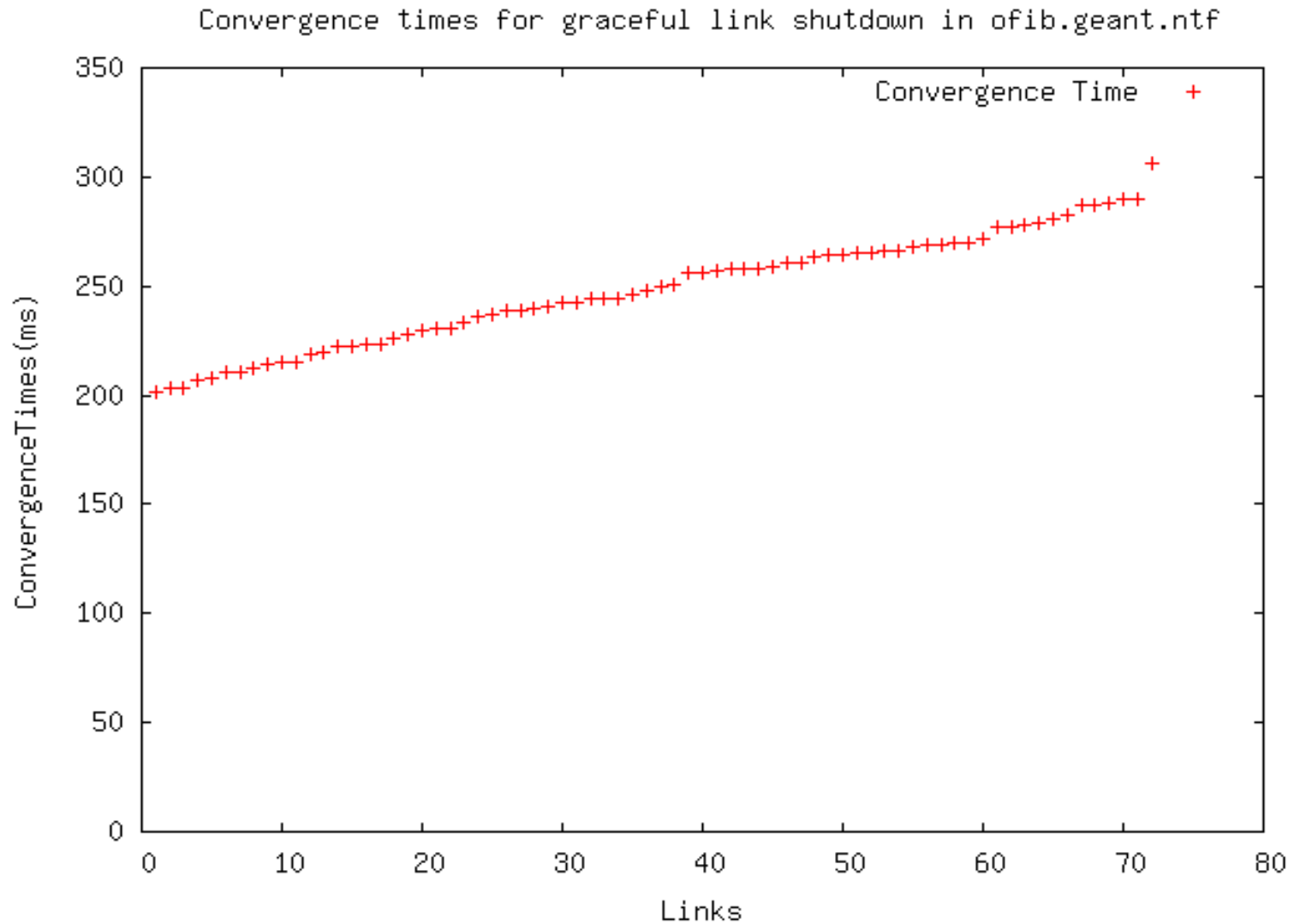
## Geant

---

- 22 routers in Europe
  - 1 access router in New-York
- 36 links (72 directed links)
- 1 asymmetrical link metric  $(X \rightarrow Y) \neq (Y \rightarrow X)$
- Few prefixes advertised by each router

# First case study

## Geant



- Convergence time with Ordered FIB similar to Normal Convergence Time



# Second topology

## A tier-1 ISP

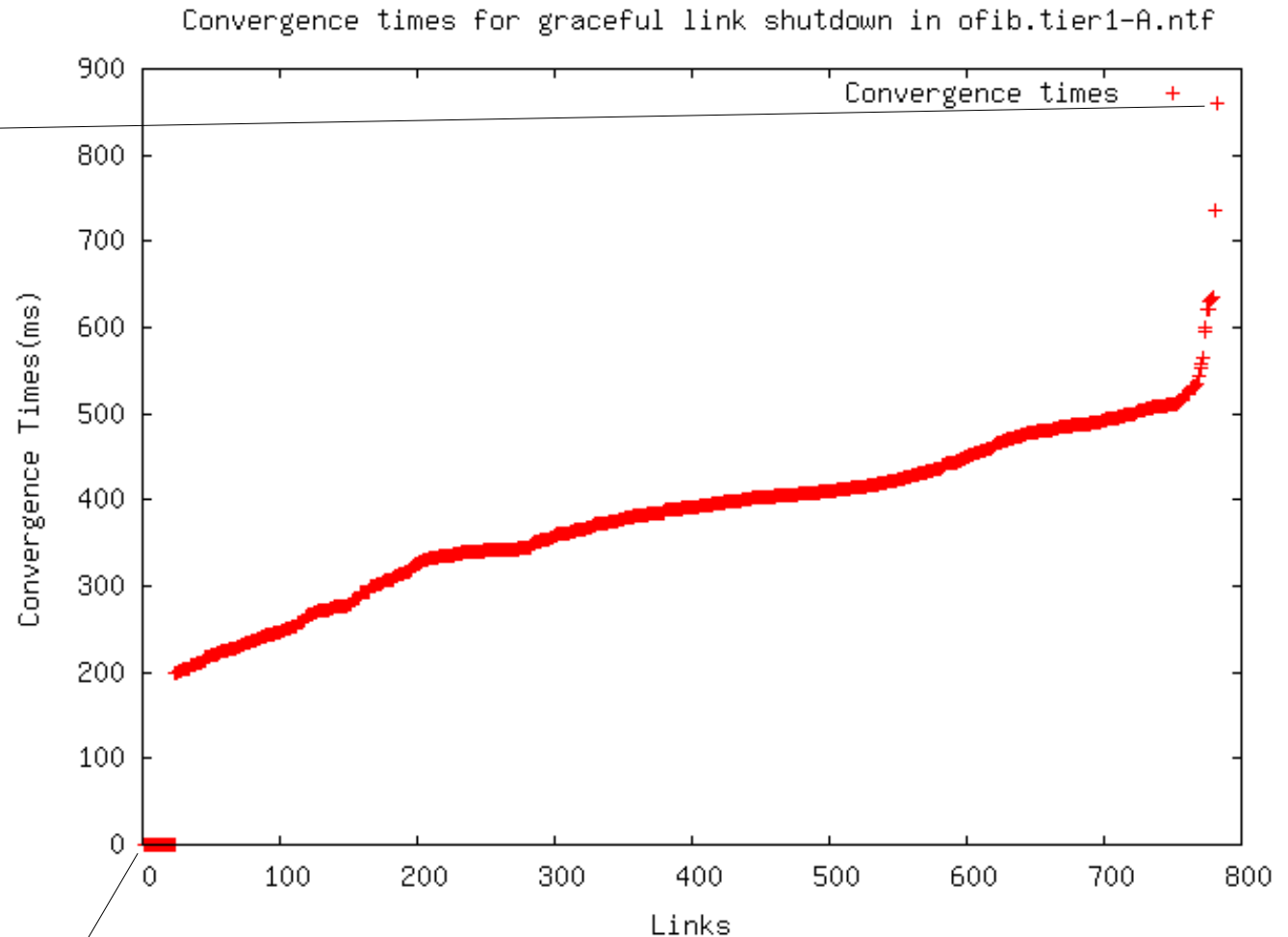
---

- 208 routers
- 391 links
- 85 asymmetrical link metrics
  - Asymmetrical link metrics are **not academic** issues
- Large number of prefixes advertised by each router

# Second topology

## A tier-1 ISP

- Worst convergence time is 861 ms
  - A branch of 4 routers in the rSPT with a full FIB update to perform, each taking more than 100 ms



- 23 directed links do not carry packets

- Convergence time with Ordered FIB similar to Normal Convergence Time

# oFIB and PLSN

## A replacement or a next-step ?

---

- Path Locking via Safe Neighbours
  - **Basic** Solution to Provide loop-free convergence
    - Does not provide 100% coverage
      - Some topology changes cannot be supported (**loops still occur**)
      - Issues with asymmetrical metrics (solution reduces the coverage)
    - A Router updates its FIB after **0, 2 or 4** seconds depending on PLSN **type** of the rerouted prefixes
- Ordered FIB updates
  - **Complete** Solution
    - Provides 100% loop-free convergence for
      - link/router/linecard manual up/down, IGP metric tuning
      - sudden failures when a local protection is provided
    - A Router updates its FIB after (rank \* max\_fib), in one shot
      - **worst-case rank** is the longest (in hops) path in the network
    - **Sub-second** convergence time can be achieved with completion messages, rank time only applies if CM are lost

# Conclusion

---

- Ordered FIB updates can provide sub-second loop-free convergence in IS-IS and OSPF networks
  - Full coverage
  - Simulations indicate that Ordered convergence is *not significantly slower* than normal IGP convergence
    - If completion messages are used
- Adoption of draft-francois-ordered-fib-01.txt as WG document
  - Requested in November 2005 for previous version
  - Comments raised on mailing list addressed by this draft
  - There is an existing implementation

# References

---

- [1] [Achieving subsecond IGP convergence in large IP networks](#)  
P. Francois, C. Filsfils, J. Evans, O. Bonaventure,  
In ACM SIGCOMM Computer Communications Review, July 2005
- [draft-bonaventure-isis-ordered-00.txt](#)
- [Avoiding transient loops during IGP convergence in IP networks](#),  
P. Francois, O. Bonaventure,  
In proceedings of IEEE INFOCOM 2005