BGP Wedgies: Bad Policy Interactions that Cannot be Debugged

Timothy G. Griffin

Intel Research,

Cambridge UK

tim.griffin@intel.com

http://www.cambridge.intel-research.net/~tgriffin/

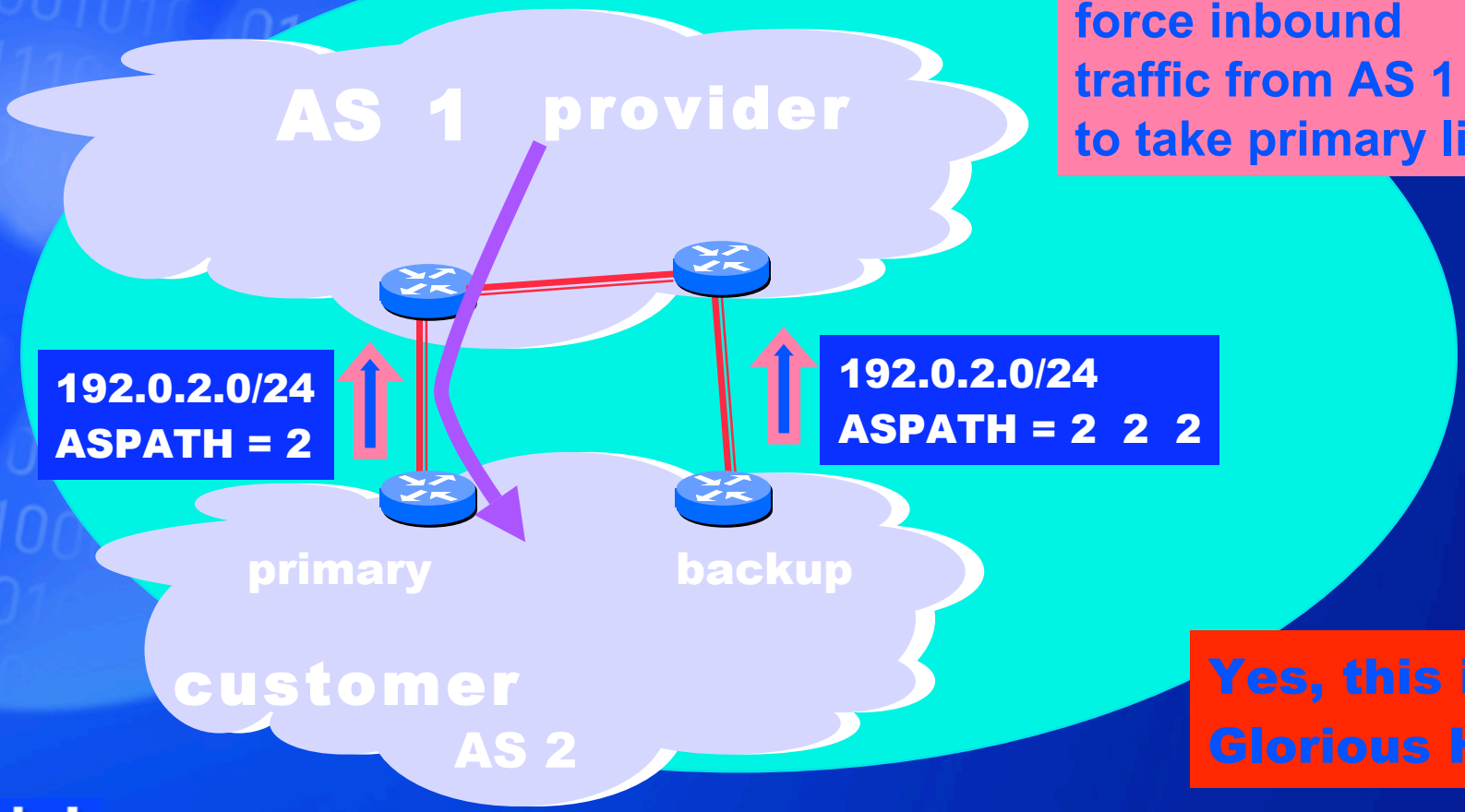
IDRWS II

Amsterdam

May 1-2, 2004

Presented by: Daniel Karrenberg

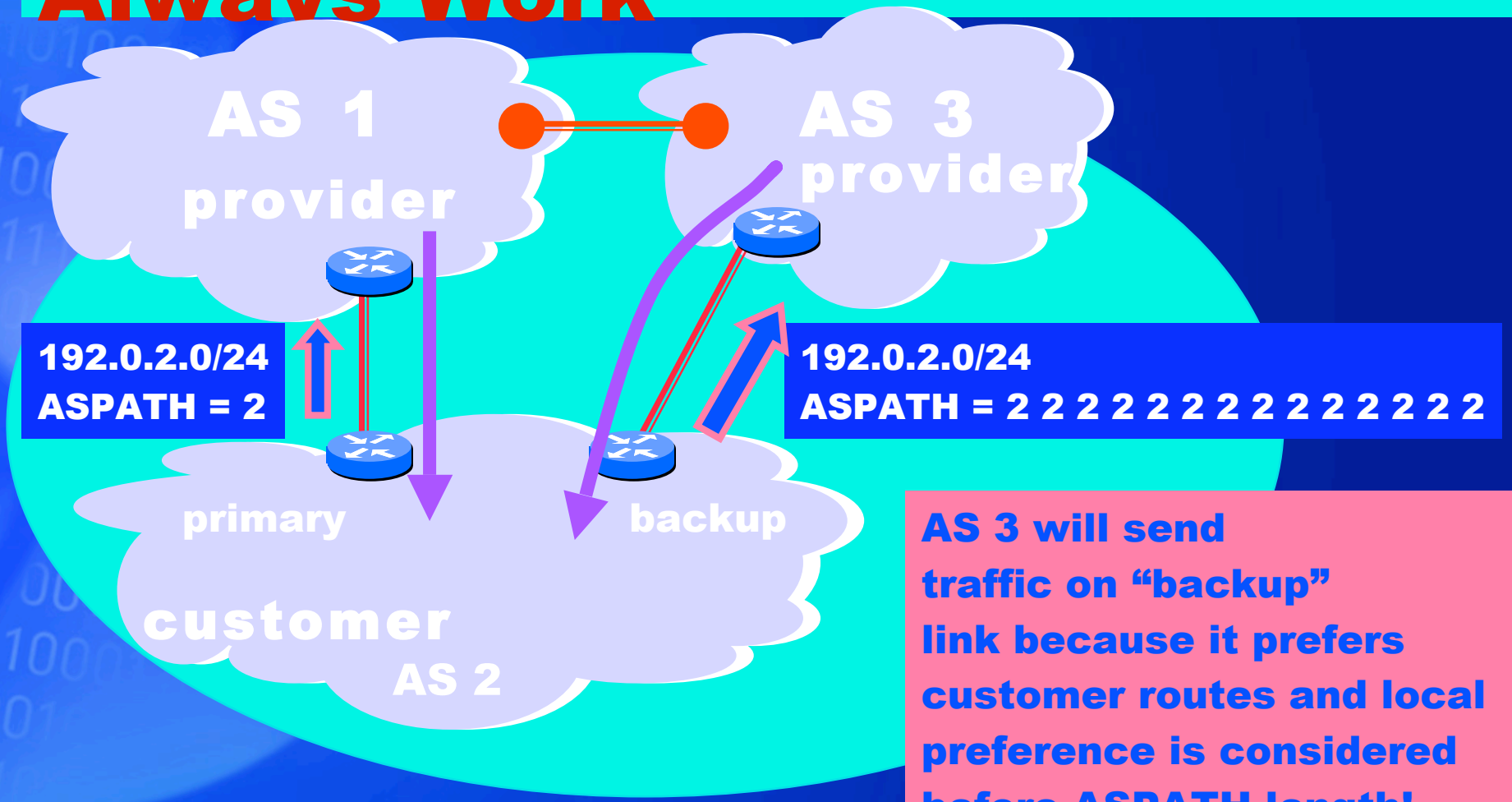
Shedding Inbound Traffic with ASPATH Prepending



Prepending will (usually) force inbound traffic from AS 1 to take primary link

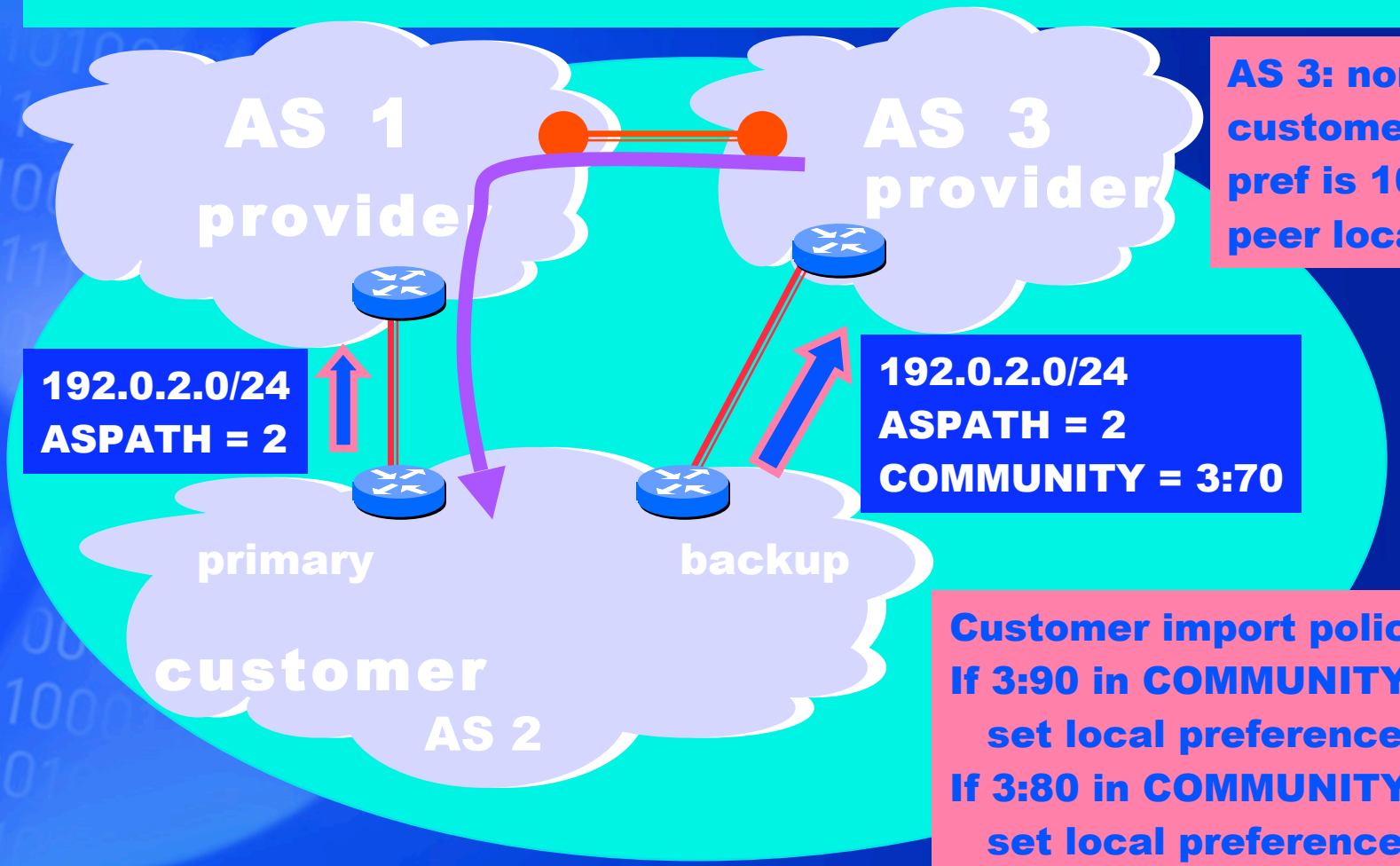
Yes, this is a Glorious Hack ...

… But Padding Does Not Always Work



AS 3 will send traffic on "backup" link because it prefers customer routes and local preference is considered before ASPATH length!

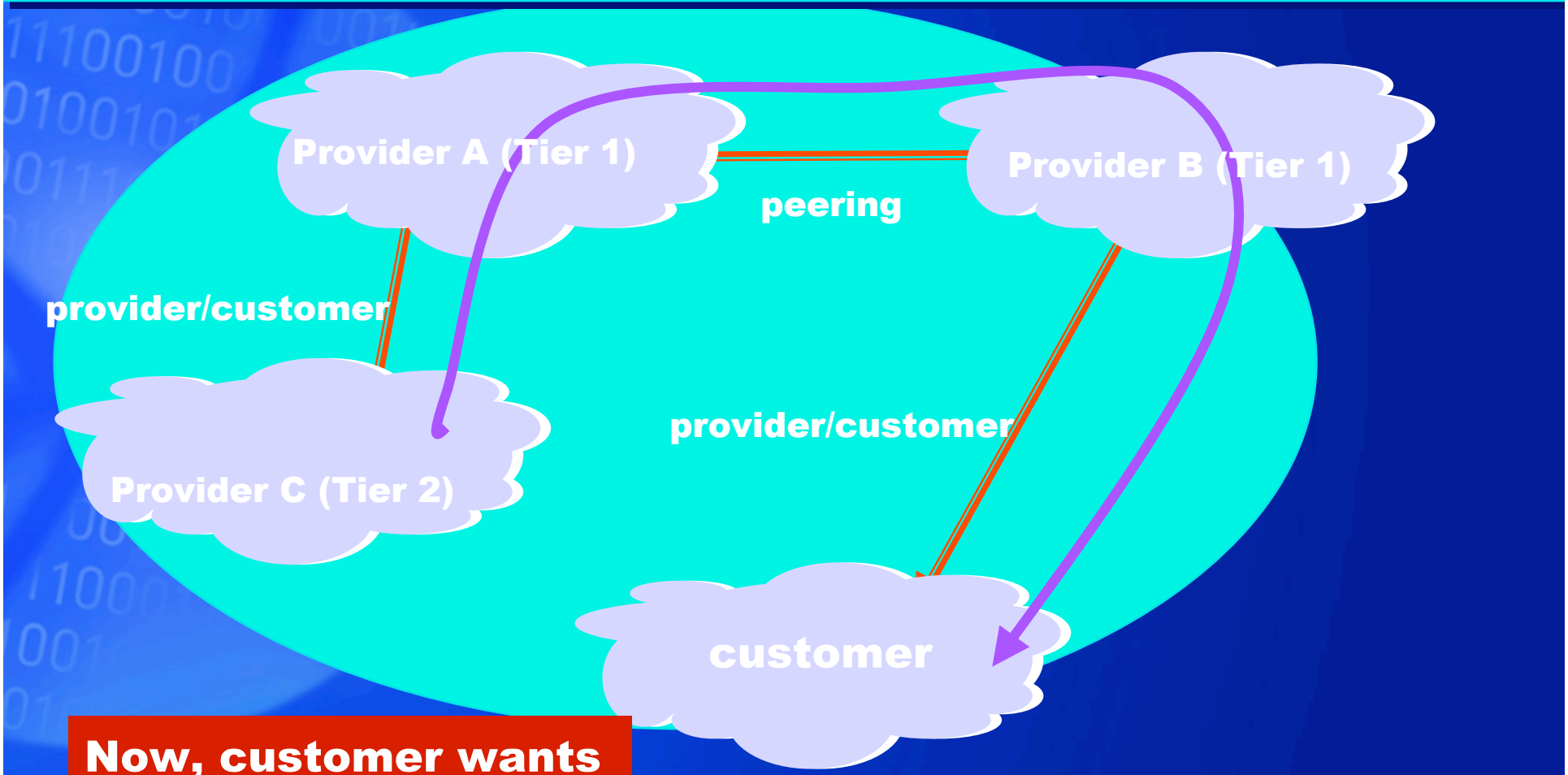
COMMUNITIES to the Rescue!



AS 3: normal customer local pref is 100, peer local pref is 90

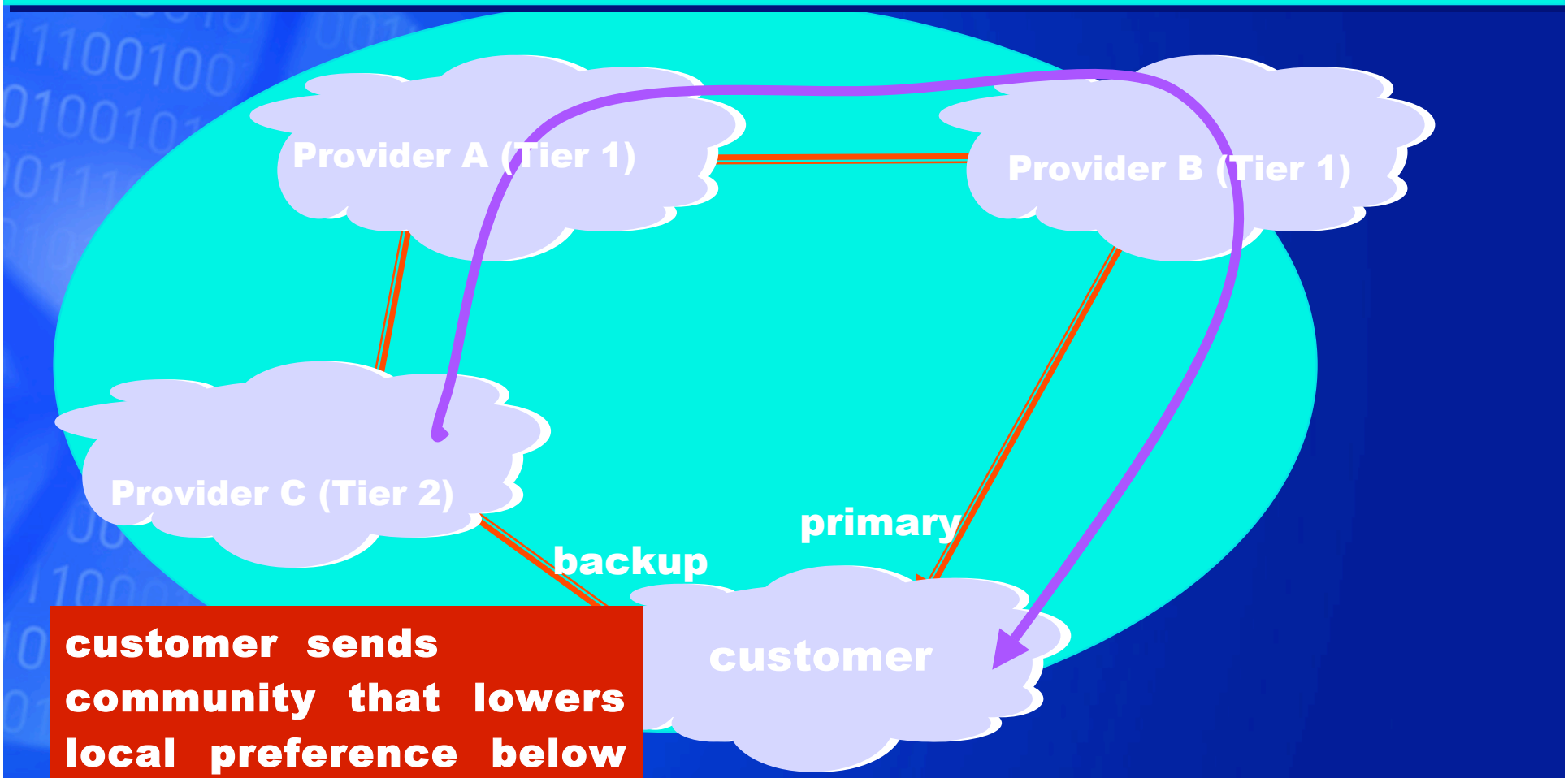
Customer import policy at AS 3:
If 3:90 in COMMUNITY then
set local preference to 90
If 3:80 in COMMUNITY then
set local preference to 80
If 3:70 in COMMUNITY then
set local preference to 70

Don’t Celebrate Just Yet….



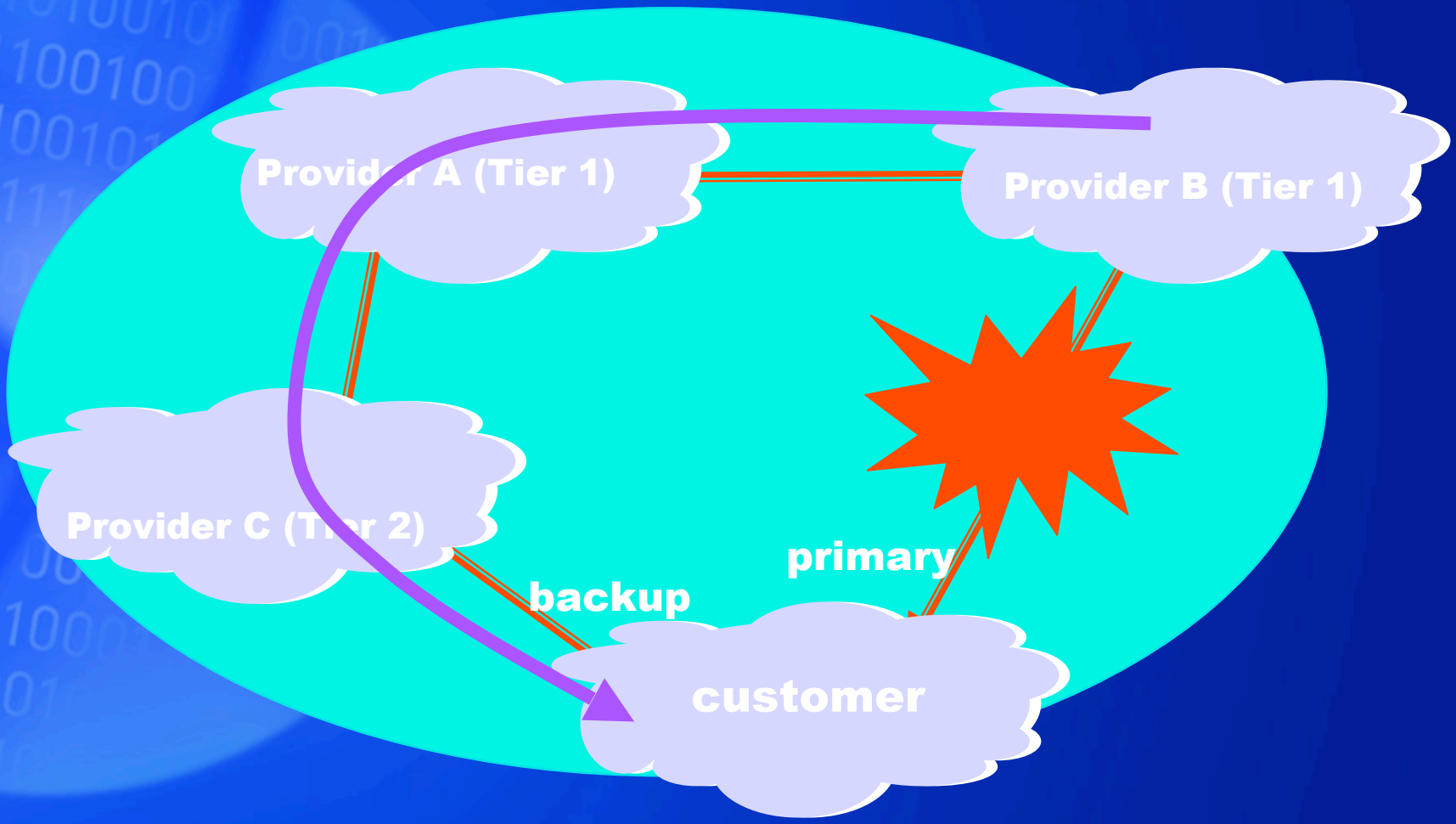
Now, customer wants a backup link to C….

Customer installs a “backup link” …



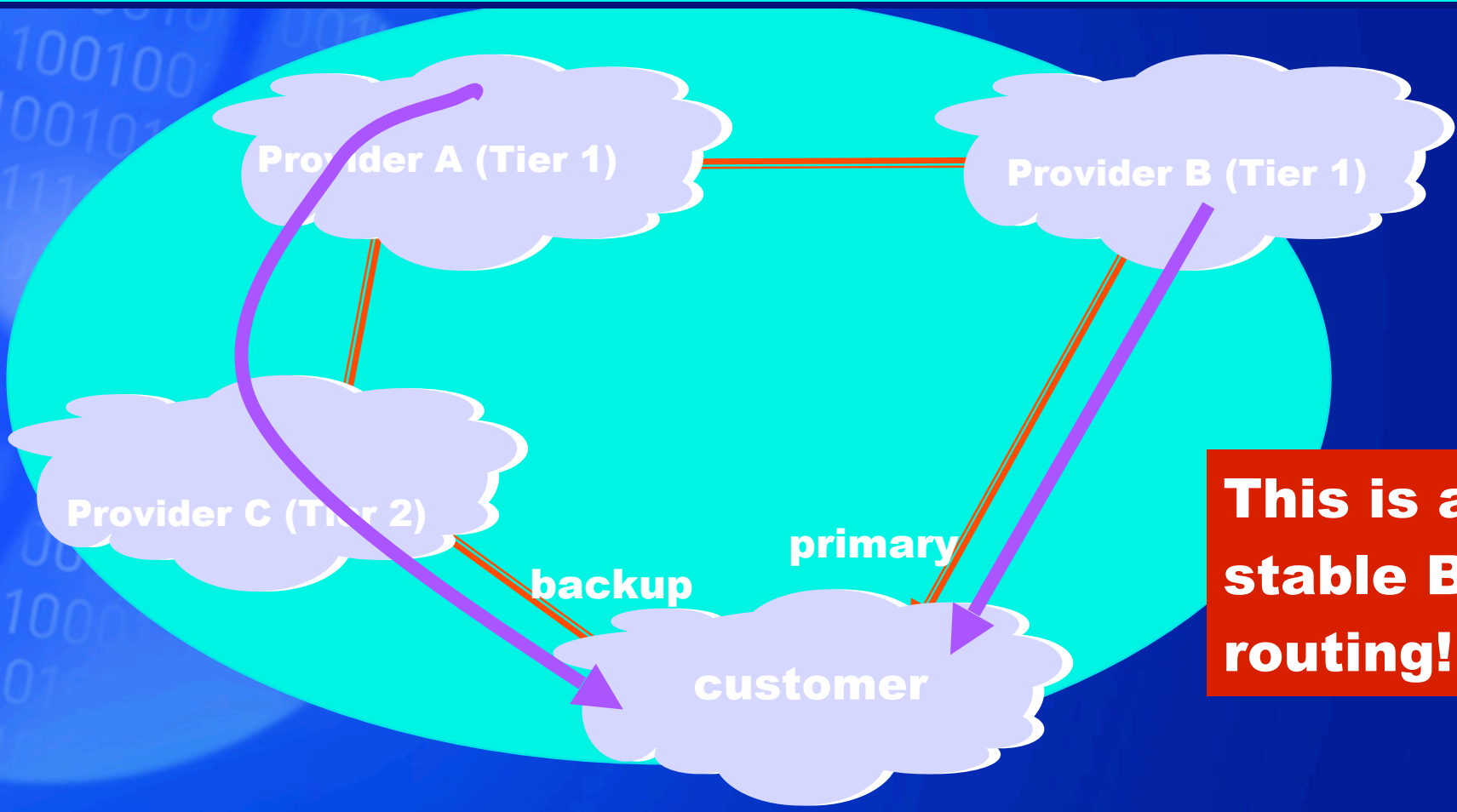
customer sends community that lowers local preference below a provider’s

Disaster Strikes!



customer is happy that backup was installed ...

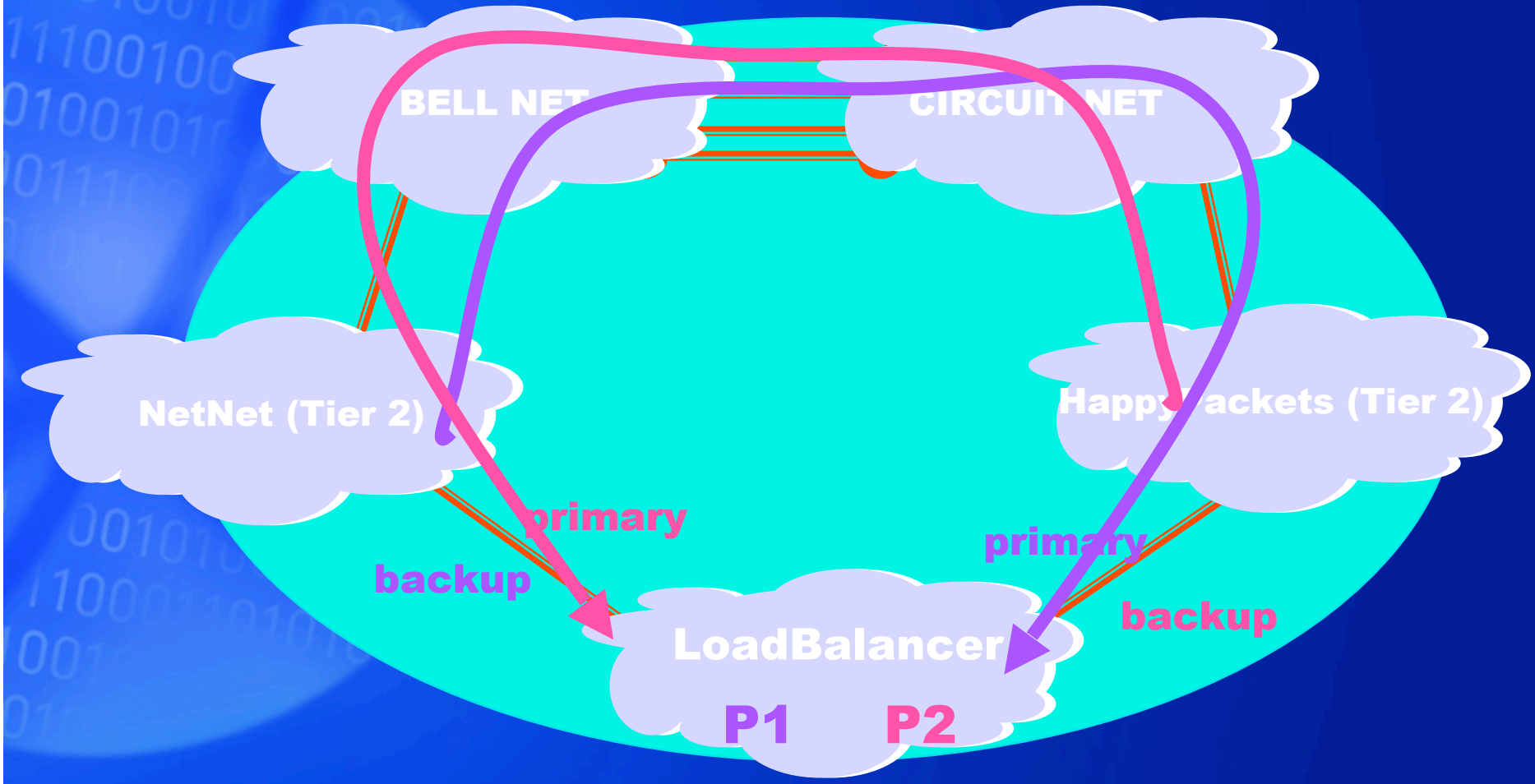
The primary link is repaired, yet routing does repair!



This is a stable BGP routing!

One "solution" --- reset BGP session on backup link!

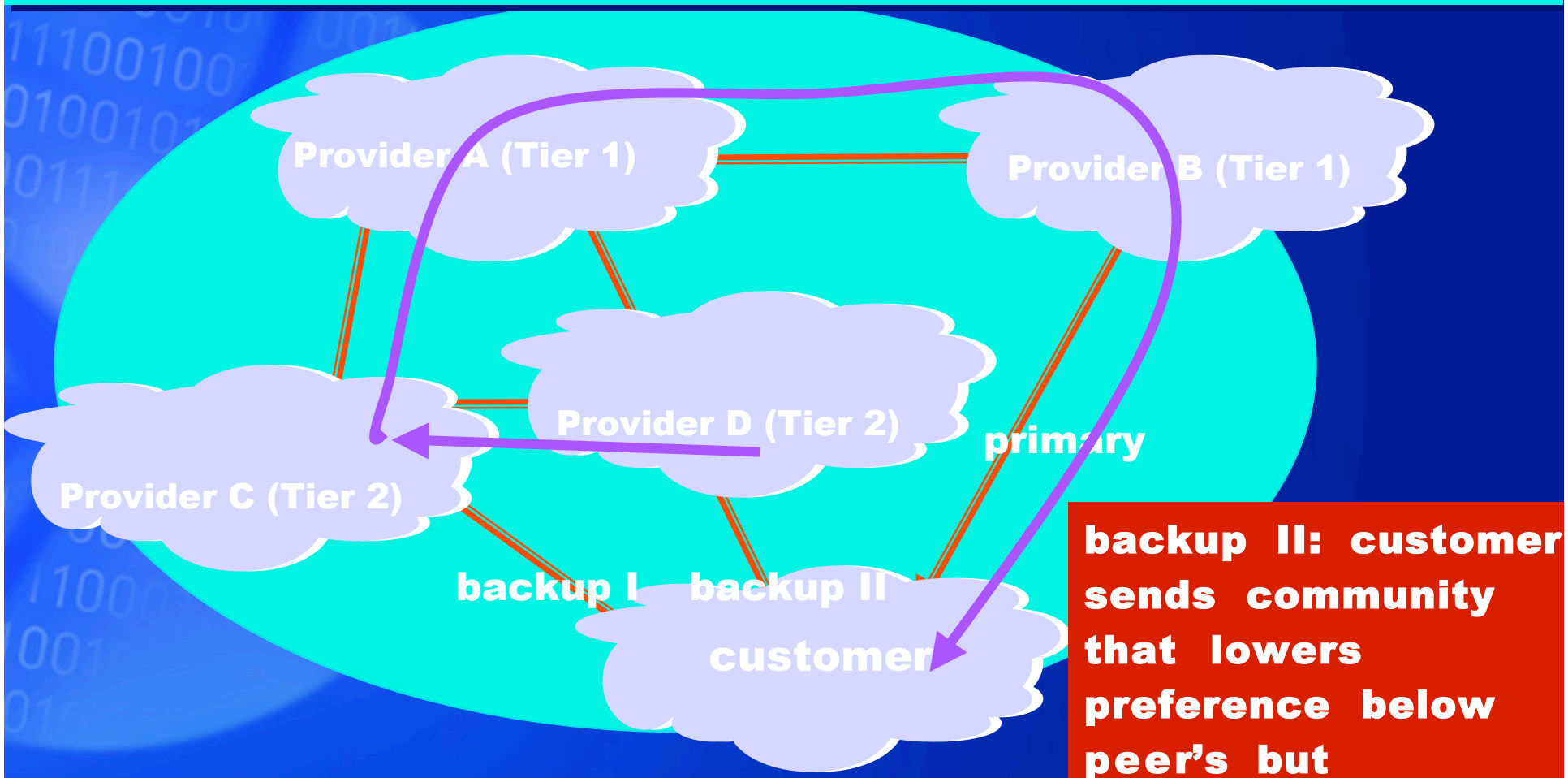
Ouch


BELL NET
CIRCUIT NET
NetNet (Tier 2)
Happy Packets (Tier 2)
primary
backup
primary
backup
LoadBalancer
P1
P2

What the heck is going on?

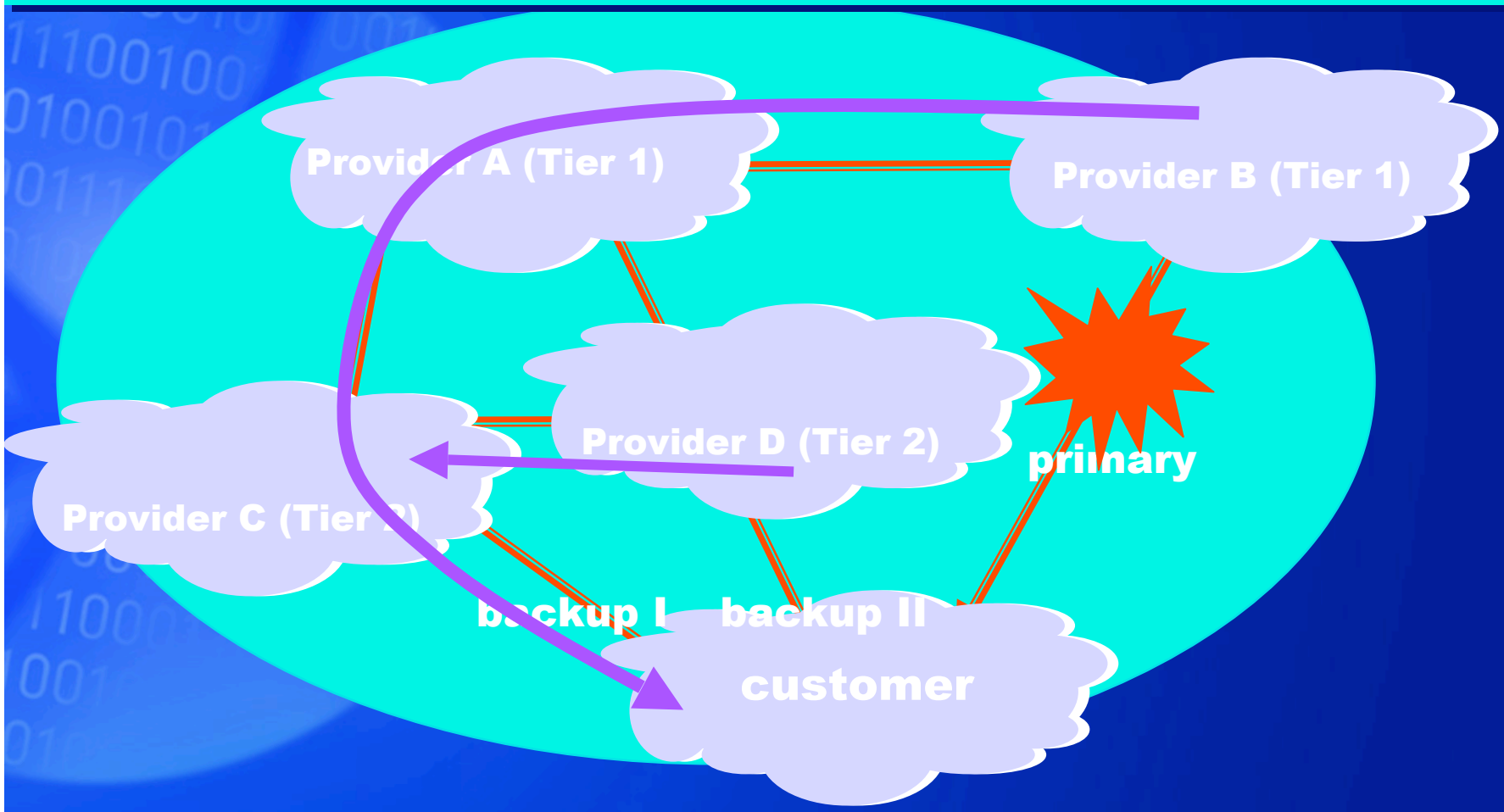
- There is no guarantee that a BGP configuration has a unique routing solution.
  - When multiple solutions exist, the (unpredictable) order of updates will determine which one is wins.
- There is no guarantee that a BGP configuration has any solution!
  - And checking configurations is NP-Complete
- Complex policies (weights, communities setting preferences, and so on) increase chances of routing anomalies.
  - … yet this is the current trend!

More fun with communities ....

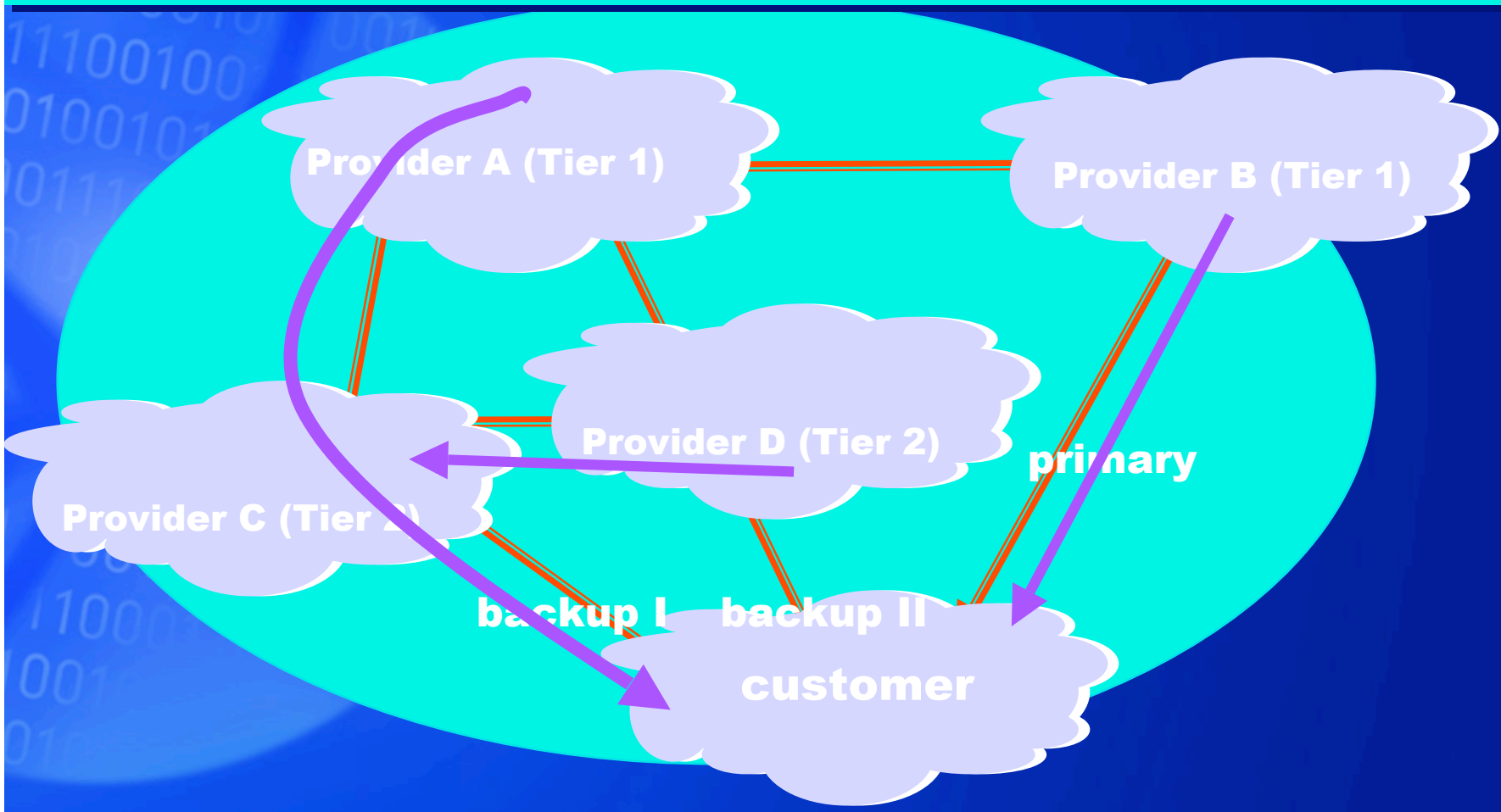


backup II: customer sends community that lowers preference below peer's but above provider's

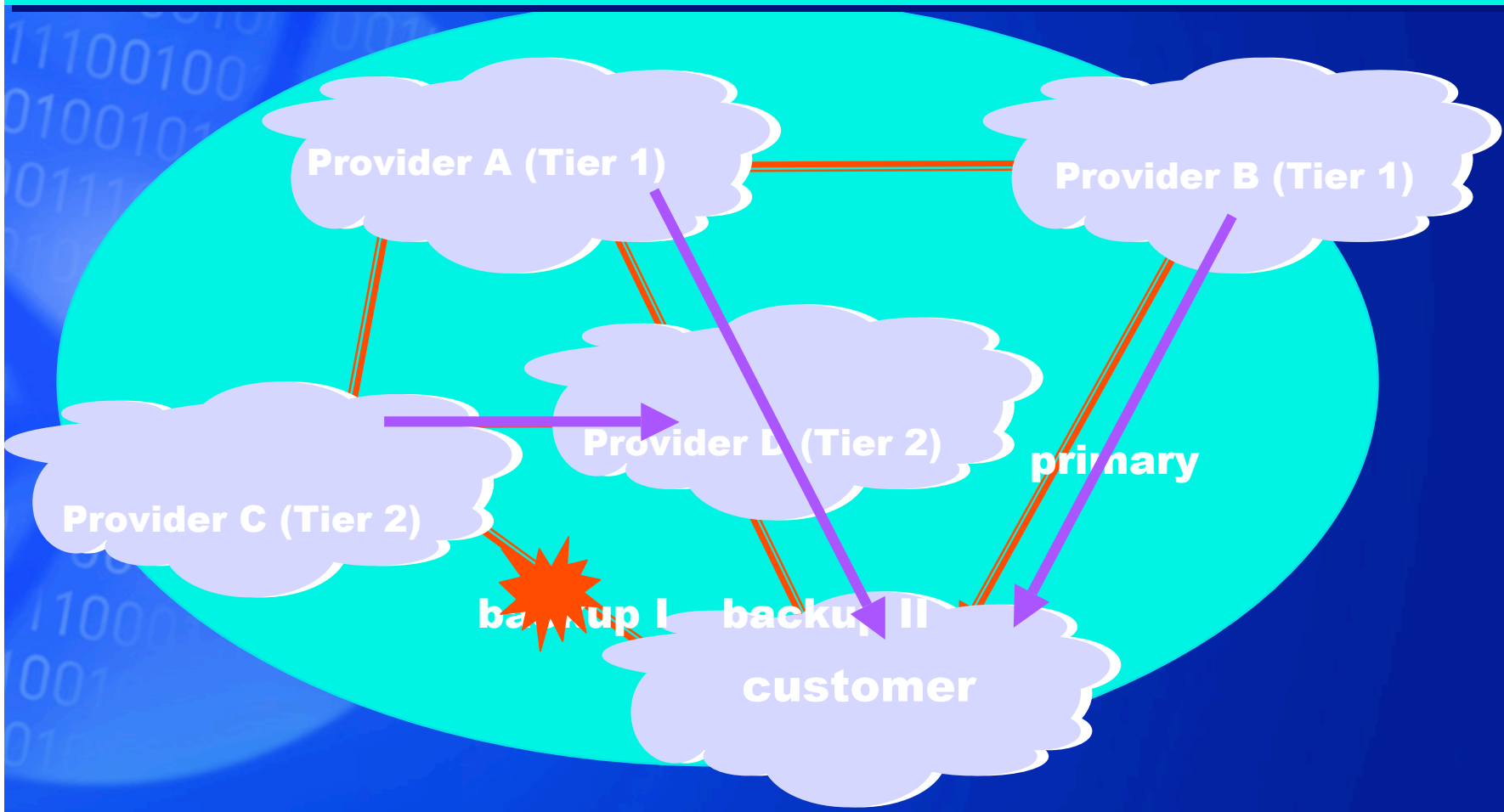
Primary goes down!



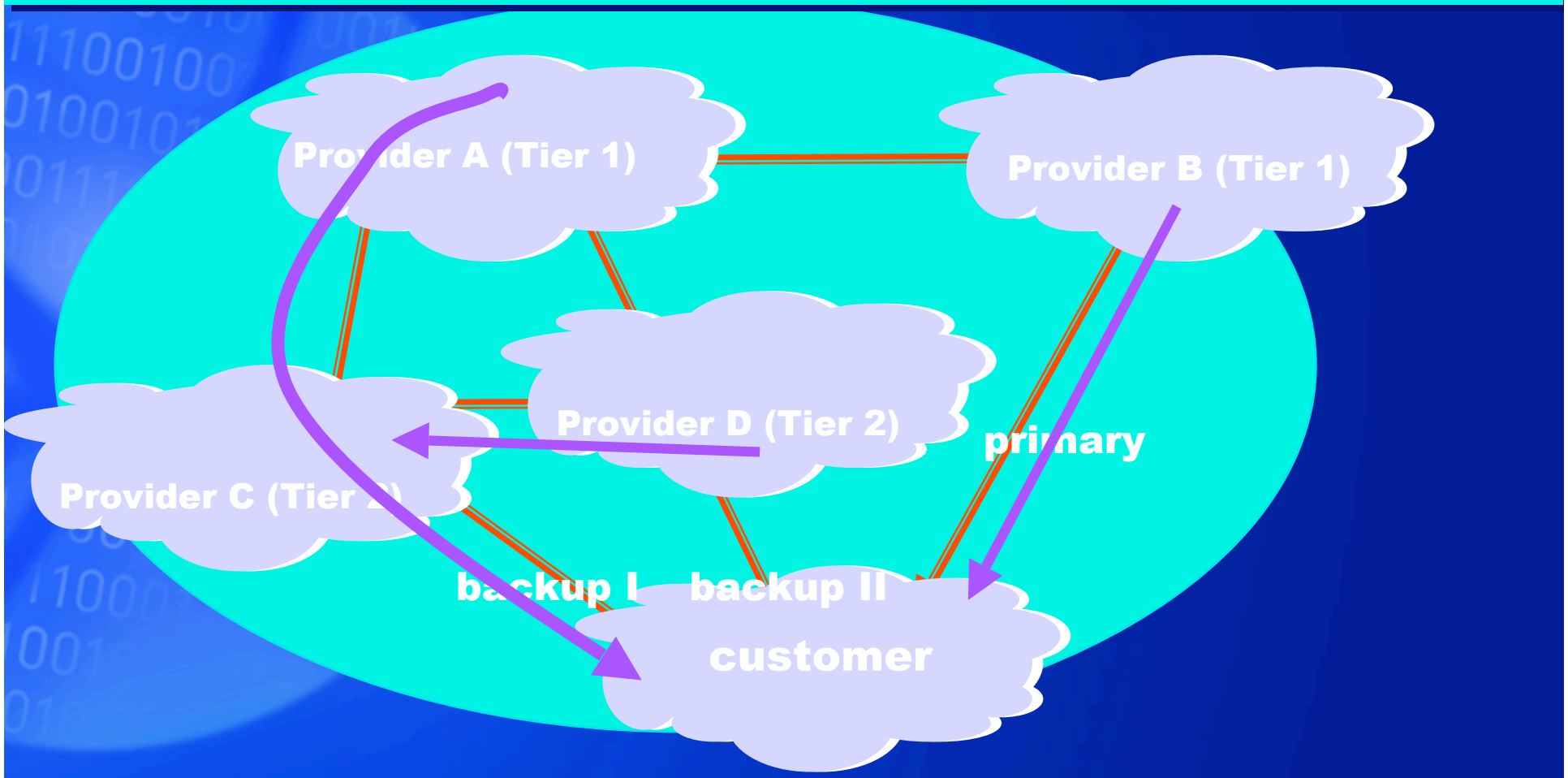
Primary repaired



Reset Backup I session? First, take it down....



Now Bring it up ...



“solution” --- reset BGP session on BOTH backup links simultaneously!

BGP Wedgie

- BBP policies make sense locally
- Sum of local policies allows multiple solutions
- Some solutions are consistent with intended policies, and some are not
- When unintended solutions are installed, no single AS has enough global knowledge to effectively debug the problem

Next?

- Need techniques for constructing policy languages.
- Design of protocols to enforce global constraints.
- Can ad-hocery be avoided?

Some References

- Persistent Route Oscillations in Inter-Domain Routing. Kannan Varadhan, Ramesh Govindan, and Deborah Estrin. Computer Networks, Jan. 2000. (Also USC Tech Report, Feb. 1996)
  - Shows that BGP is not guaranteed to converge
- An Architecture for Stable, Analyzable Internet Routing. Ramesh Govindan, Cengiz Alaettinoglu, George Eddy, David Kessens, Satish Kumar, and WeeSan Lee. IEEE Network Magazine, Jan-Feb 1999.
  - Use RPSL to specify policies. Store them in registries. Use registry for conguration generation and analysis.
- An Analysis of BGP Convergence Properties. Timothy G. Griffin, Gordon Wilfong. SIGCOMM 1999
  - Model BGP, shows static analysis of divergence in policies is NP complete
- Policy Disputes in Path Vector Protocols. Timothy G. Griffin, F. Bruce Shepherd, Gordon Wilfong. ICNP 1999
  - Define Stable Paths Problem and develop sufficient condition for "sanity"
- A Safe Path Vector Protocol. Timothy G. Griffin, Gordon Wilfong. INFOCOM 2001
  - Dynamic solution for SPVP based on histories
- Stable Internet Routing without Global Coordination. Lixin Gao, Jennifer Rexford. SIGMETRICS 2000
  - Show that if certain guidelines are followed, then all is well.
- Inherently safe backup routing with BGP. Lixin Gao, Timothy G. Griffin, Jennifer Rexford. INFOCOM 2001
  - Use SPP to study complex backup policies
- On the Correctness of IBGP Configurations. Griffin and Wilfong.SIGCOMM 2002.
- An Analysis of the MED oscillation Problem. Griffin and Wilfong. ICNP 2002.
- Design Principles of Policy Languages for Path Vector Protocols. Timothy G. Griffin, Aaron D. Jaggard (Department of Mathematics, University of Pennsylvania), Vijay Ramachandran (Department of Computer Science, Yale University). SIGCOMM 2003.
- Network Routing with Path Vector Protocols: Theory and Applications. Joao Luis Sobrinho SIGCOMM 2003

Questions?

Again: The work is Tim's

(Un)clearness of Presentation: Daniel