BGP Aggregation & The Deaggregation Report

Philip Smith

SANOG 12
6th-14th August 2008
Kathmandu

Route Aggregation Recommendations

- LINX attempted aggregation policy for members
  - It failed even though most members voted for policy
- RIPE Routing Working Group work item from early 2006
  - Based on early LINX concept
  - Authored by Philip Smith, Mike Hughes (LINX) and Rob Evans (UKERNA)

Route Aggregation Recommendations

- RIPE Document — RIPE-399
  - http://www.ripe.net/ripe/docs/ripe-399.html
- Discusses:
  - History of aggregation
  - Causes of de-aggregation
  - Impacts on global routing system
  - Available Solutions
  - Recommendations for ISPs

History:

- **Classful to classless migration**
  - Clean-up efforts in 192/8
- **CIDR Report**
  - Started by Tony Bates to encourage adoption of CIDR & aggregation
  - Mostly ignored through late 90s
  - Now part of extensive BGP table analysis by Geoff Huston
- **Introduction of Regional Internet Registry system and PA address space**

Deaggregation: Claimed causes (1):

- **Routing System Security**
  - "Announcing /24s means that no one else can DOS the network"

- **Reduction of DOS attacks & miscreant activities**
  - "Announcing only address space in use as rest attracts 'noise'"

- **Commercial Reasons**
  - "Mind your own business"

Deaggregation: Claimed causes (2):

- Leakage of iBGP outside of local AS
    - eBGP is NOT iBGP – how many ISPs know this?
- Traffic Engineering for Multihoming
    - Spraying out /24s hoping it will work
    - Rather than do any **real engineering**
- Legacy Assignments
    - "All those pre-RIR assignments are to blame"
    - In reality it is both RIR and legacy assignments

Impacts (1):

- **Router memory**
  - Shortens router life time as vendors underestimate memory growth requirements
  - Depreciation life-cycle shortened
  - Increased costs for ISP and customers
- **Router processing power**
  - Processors are underpowered as vendors underestimate CPU requirement
  - Depreciation life-cycle shortened
  - Increased costs for ISP and customers

Impacts (2):

- **Routing System convergence**
  - Larger routing table → slowed convergence
  - Can be improved by faster control plane processors — see earlier
- **Network Performance & Stability**
  - Slowed convergence → slowed recovery from failure
  - Slowed recovery → longer downtime
  - Longer downtime → unhappy customers

Solutions (1):

- **CIDR Report**
  - Global aggregation efforts
  - Running since 1994
- **Routing Table Report**
  - Per RIR region aggregation efforts
  - Running since 1999
- **Filtering recommendations**
  - BCP38, training, tutorials, Team Cymru,…
- **"CIDR Police"**

Solutions (2):

- BGP Features:
  - NO_EXPORT Community
  - NOPEER Community
    - RFC3765 — but no one has implemented it
  - AS_PATHLIMIT attribute
    - Still working through IETF IDR Working Group
  - Provider Specific Communities
    - Some ISPs use them; most do not

RIPE-399 Recommendations:

- Announcement of initial allocation as a single entity
- Subsequent allocations aggregated if they are contiguous and bit-wise aligned
- Prudent subdivision of aggregates for Multihoming
- Use BGP enhancements already discussed
- (All of this applies to IPv6 too)

Looking at Deaggregation

- CIDR Report
  - www.cidr-report.org
  - Encourages aggregation following CIDRisation of Internet
  - Today: extensive suite of reports and tools covering state of BGP table
- Routing Report
  - BGP table status on per RIR basis
  - Original CIDR Report and a whole lot more

Deaggregation Factor

- **Routing Report**
  - One summary takes BGP table and aggregates prefixes by origin AS
    - Called "Max Aggregation" in report
  - Global and per RIR basis
    - http://thyme.apnic.net/current/
- **New Deaggregation Factor:**
  - Measure of Routing Table size/Aggregated Size
  - Global value has been increasing slowly and steadily since "records began"

August 2008

Total Prefixes

- Global BGP Table
  - 266k prefixes

- Europe & Middle East
  - 57k prefixes
- North America
  - 122k prefixes
- Asia & Pacific
  - 61k prefixes
- Africa
  - 4k prefixes
- Latin America & Caribbean
  - 21k prefixes

Deaggregation Factor

- Global Average
  - 2.05

- Europe & Middle East
  - 1.64
- North America
  - 1.87
- Asia & Pacific
  - 2.69
- Africa
  - 3.29
- Latin America & Caribbean
  - 3.96

Deaggregation: RIR Regions vs Global

Ratio

Days since 23rd February 1999
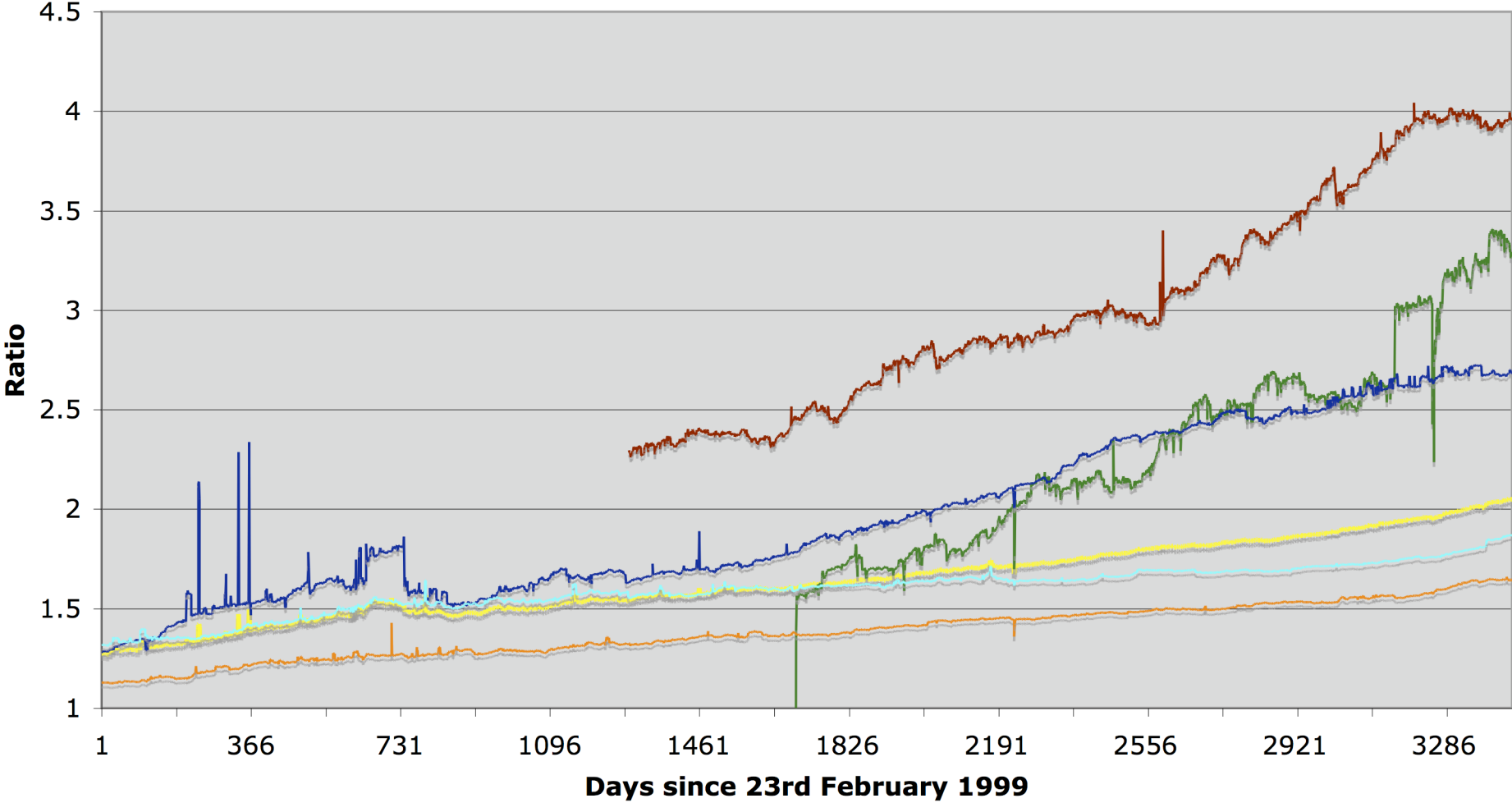
Global — AfriNIC — APNIC — ARIN — LACNIC — RIPE

```
Africa Aggregation Savings Summary

 ASN    No of Nets       Savings    Description
24863      495            467       LINKdotNET AS number
20858      402            399       EgyNet
 6713      143            132       Itissalat Al-MAGHRIB
33783      136            123       EEPAD TISP TELECOM & INTERNET
33776      114            111       Starcomms Nigeria Limited
 5536      120            103       Internet Egypt Network
29571      107             99       Ci Telecom Autonomous system
 2018      211             81       Tertiary Education Network
15475       67             62       Nile Online
24835       63             57       RAYA Telecom - Egypt
15706       61             57       Sudatel Internet Exchange Aut
 5713      129             55       Telkom SA Ltd
29975       62             47       Vodacom
 3741      269             45       The Internet Solution
12455       45             42       Jambonet Autonomous system
 8094       42             39       PUKNET
23889       49             33       MAURITIUS TELECOM
21152       32             31       AS for the uplinks of Soficom
33774       55             28       AS Number for Telecom Algeria
 8524       32             27       AUCEGYPT Autonomous System
```

http://thyme.apnic.net/current/data-CIDRnet-AFRINIC

Asia & Pacific Aggregation Savings Summary

ASN	No of Nets	Savings	Description
4755	1691	1513	Videsh Sanchar Nigam Ltd. Aut
17488	1290	1194	Hathway IP Over Cable Interne
9583	1174	682	Sify Limited
18101	692	659	Reliance Infocom Ltd Internet
4780	713	652	Digital United Inc.
9498	674	619	BHARTI BT INTERNET LTD.
9829	598	586	BSNL National Internet Backbo
4766	858	507	Korea Telecom (KIX)
4808	618	481	CNCGROUP IP network: China169
4134	835	480	CHINANET-BACKBONE
7545	539	468	TPG Internet Pty Ltd
17676	524	462	Softbank BB Corp.
17974	479	462	PT TELEKOMUNIKASI INDONESIA
9443	501	423	Primus Telecommunications
24560	557	410	Bharti Airtel Ltd.
10091	344	333	SCV Broadband Access Provider
4802	494	330	Wantree Development
4668	334	327	LG-EDS Systems Inc.
23966	328	310	Dancom Pakistan (PVT) Limited
7552	302	298	Vietel Corporation

http://thyme.apnic.net/current/data-CIDRnet-APNIC

North America Aggregation Savings Summary

ASN	No of Nets	Savings	Description
209	2973	2349	Qwest
1785	1453	1351	AppliedTheory Corporation
11492	1215	1193	Cable One
4323	1489	1109	Time Warner Telecom
18566	1045	1035	Covad Communications
6298	1778	1010	Cox Communicatons
22773	974	911	Cox Communications, Inc.
6478	1016	843	AT&T Worldnet Services
19262	935	768	Verizon Global Networks
6517	722	674	Yipes Communications, Inc.
2386	1560	667	AT&T Data Communications Serv
5668	692	652	CenturyTel Internet Holdings,
3356	975	554	Level 3 Communications, LLC
20115	1116	544	Charter Communications
855	587	534	Canadian Research Network
6197	952	477	BellSouth Network Solutions,
7011	973	462	Citizens Utilities
33588	456	430	Bresnan Communications, LLC.
7018	1411	422	AT&T WorldNet Services
8103	618	384	Florida Department of Managem

http://thyme.apnic.net/current/data-CIDRnet-ARIN

Latin America Aggregation Savings Summary

ASN	No of Nets	Savings	Description
8151	1273	1046	UniNet S.A. de C.V.
11830	604	595	Instituto Costarricense de El
22047	565	551	VTR PUNTO NET S.A.
16814	426	416	NSS, S.A.
7303	469	404	Telecom Argentina Stet-France
14117	375	366	Telefonica del Sur S.A.
6471	411	363	ENTEL CHILE S.A.
11172	410	340	Servicios Alestra S.A de C.V
10620	404	339	TVCABLE BOGOTA
10481	310	301	Prima S.A.
28573	303	274	NET Servicos de Comunicao S.A
20299	335	237	NEWCOM AMERICAS
14259	296	235	GTD Internet S.A.
7738	252	226	Telecomunicacoes da Bahia S.A
14522	194	186	SatNet S.A.
19169	205	184	Telconet
23216	243	183	RAMtelecom Telecomunicaciones
8163	187	174	METROTEL REDES S.A.
21826	205	164	INTERCABLE
6458	173	157	GUATEL

http://thyme.apnic.net/current/data-CIDRnet-LACNIC

EU & Middle East Aggregation Savings Summary

ASN	No of Nets	Savings	Description
8452	357	346	TEDATA
8866	316	295	Bulgarian Telecommunication C
5462	296	269	Telewest Broadband
3215	371	265	France Telecom Transpac
8551	287	249	Bezeq International
9155	251	237	QualityNet AS number
35141	241	235	Megalan Autonomous system of
12479	231	225	Uni2 Autonomous System
9121	251	224	TTnet Autonomous System
29357	217	213	WATANIYA TELECOM
3352	244	202	Ibernet, Internet Access Netw
9198	208	197	Kazakhtelecom Data Network Ad
3269	244	171	TELECOM ITALIA
8877	178	171	BOL.BG Autonomous System
6830	188	145	UPC Distribution Services
3300	232	133	AUCS Communications Services
29314	150	129	Telewizja Kablowa Dami Sp. z
5384	136	125	Emirates Internet
9051	152	118	INCONET Autonomous System
1267	153	117	Infostrada S.p.A.

http://thyme.apnic.net/current/data-CIDRnet-RIPE

Observations

- Range of operational "practices" between RIR regions
  - Deaggregation by newer ISPs & developing regions is growing rapidly
  - Is harming the **entire** Internet
- RIPE-399 is only a recommendation
  - Hopefully all the RIRs will include pointers with each address allocation
  - Hopefully more ISPs will pay attention to it
  - Training is there — most ISPs choose to ignore it

Conclusion

- Make RIPE-399 your BGP good practice document