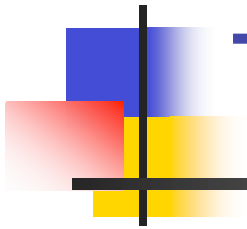
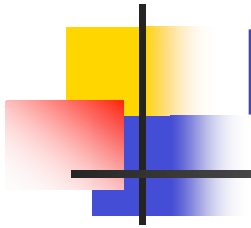


BGP Aggregation & The Deaggregation Report

Philip Smith

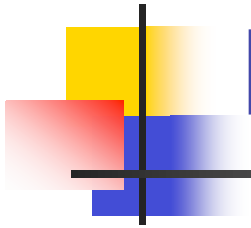
SANOG 13
Lahore
15-22 January 2009

Route Aggregation Recommendations

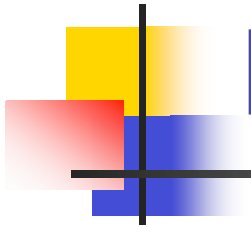


- LINX attempted aggregation policy for members
  - It failed even though most members voted for policy
- RIPE Routing Working Group work item from early 2006
  - Based on early LINX concept
  - Authored by Philip Smith, Mike Hughes (LINX) and Rob Evans (UKERNA)

Route Aggregation Recommendations

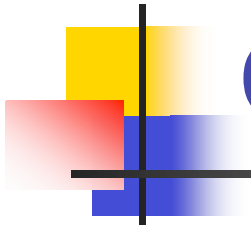


- RIPE Document — RIPE-399
  - http://www.ripe.net/ripe/docs/ripe-399.html
- Discusses:
  - History of aggregation
  - Causes of de-aggregation
  - Impacts on global routing system
  - Available Solutions
  - Recommendations for ISPs



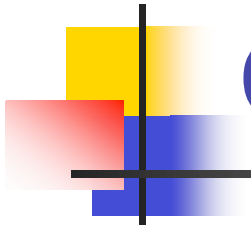
History:

- 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



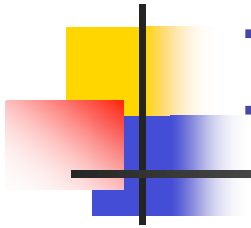
Deaggregation: Claimed causes (1):

- 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"



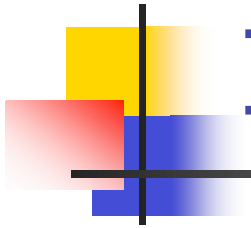
Deaggregation: Claimed causes (2):

- 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



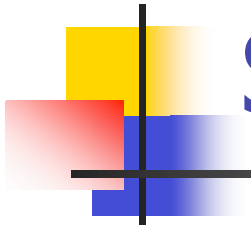
Impacts (1):

- 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



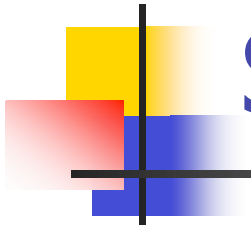
Impacts (2):

- 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



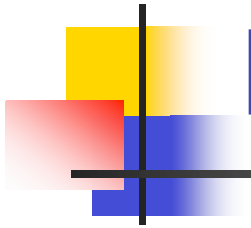
Solutions (1):

- 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”



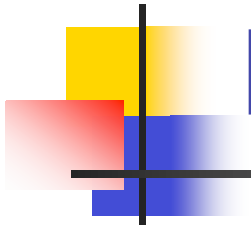
Solutions (2):

- 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



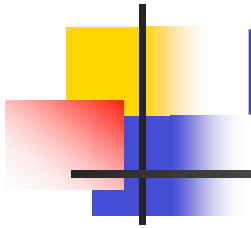
RIPE-399 Recommendations:

- 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)



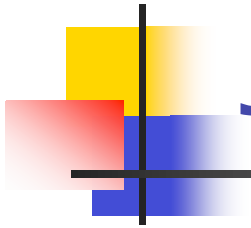
Looking at Deaggregation

- 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"



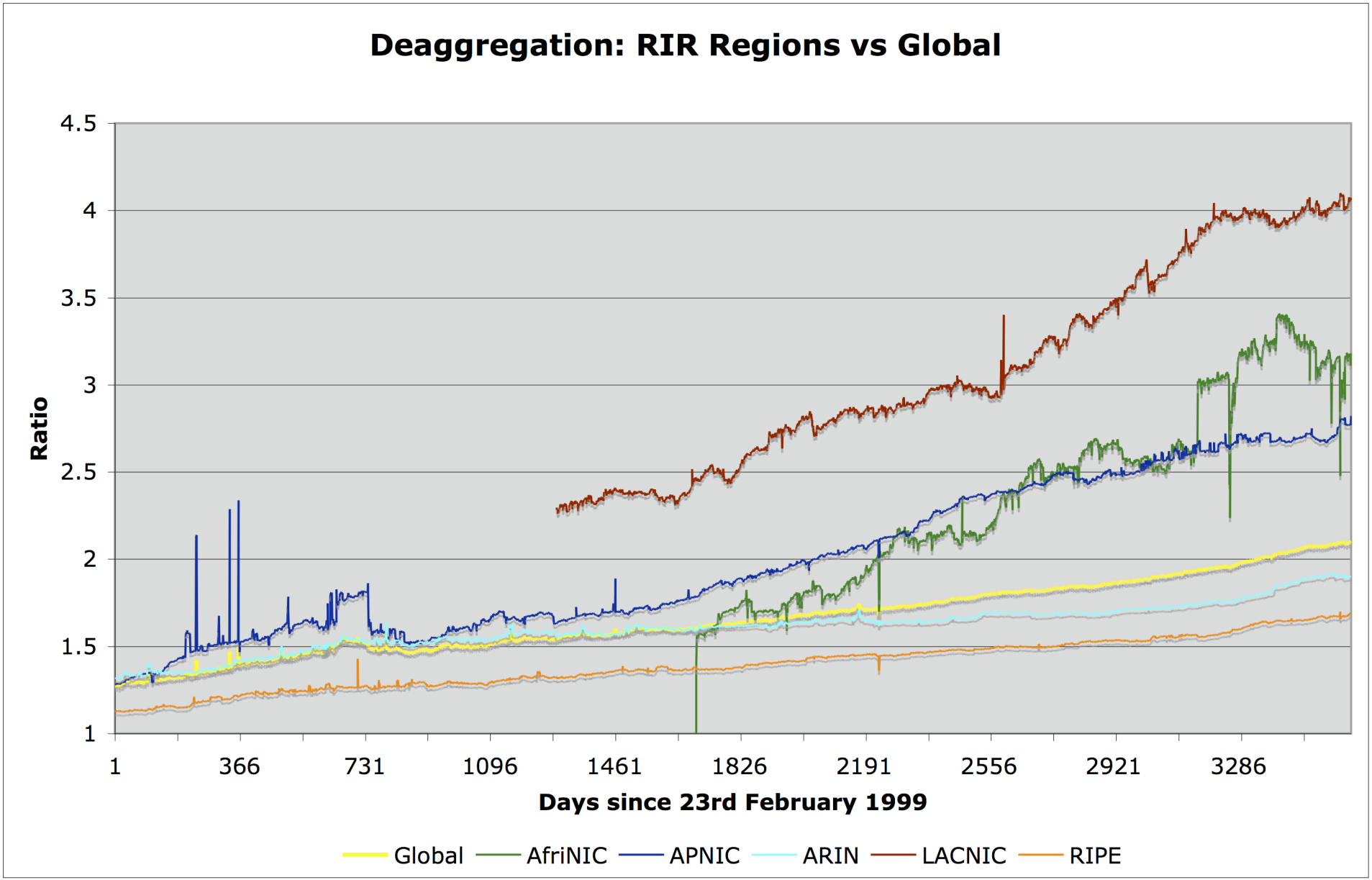
January 2009

Total Prefixes

- Global BGP Table
  - 278k prefixes
- Europe & Middle East
  - 62k prefixes
- North America
  - 124k prefixes
- Asia & Pacific
  - 64k prefixes
- Africa
  - 4k prefixes
- Latin America & Caribbean
  - 23k prefixes

Deaggregation Factor

- Global Average
  - 2.10
- Europe & Middle East
  - 1.69
- North America
  - 1.90
- Asia & Pacific
  - 2.82
- Africa
  - 3.18
- Latin America & Caribbean
  - 4.06


Deaggregation: RIR Regions vs Global
Ratio
4.5
4
3.5
3
2.5
2
1.5
1
1
366
731
1096
1461
1826
2191
2556
2921
3286
Days since 23rd February 1999
Global
AfriNIC
APNIC
ARIN
LACNIC
RIPE

Africa Aggregation Savings Summary

ASN	No of Nets	Savings	Description
24863	593	553	LINKdotNET AS number
20858	269	266	This AS will be used to conne
33783	150	142	EEPAD TISP TELECOM & INTERNET
6713	144	133	Itissalat Al-MAGHRIB
29571	128	120	Ci Telecom Autonomous system
33776	113	110	Starcomms Nigeria Limited
5536	120	103	Internet Egypt Network
2018	240	99	Tertiary Education Network
24835	101	92	RAYA Telecom - Egypt
15475	65	60	Nile Online
20928	61	59	Noor Advanced Technologies AS
15706	62	58	Sudatel Internet Exchange Aut
5713	117	51	Telkom SA Ltd
29975	62	47	Vodacom
3741	269	42	The Internet Solution
8094	42	39	PUKNET
12455	39	36	Jambonet Autonomous system
21152	32	31	AS for the uplinks of Soficom
20484	33	30	Yalla Online Autonomous Syste
36915	32	30	Africa Online Kenya

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

Asia & Pacific Aggregation Savings Summary

ASN	No of Nets	Savings	Description
17488	1492	1395	Hathway IP Over Cable Interne
4766	1666	1290	Korea Telecom (KIX)
4755	1186	1020	TATA Communications formerly
9583	1122	1011	Sify Limited
18101	778	753	Reliance Infocom Ltd Internet
4780	719	654	Digital United Inc.
9498	678	626	BHARTI BT INTERNET LTD.
9829	623	609	BSNL National Internet Backbo
7545	673	591	TPG Internet Pty Ltd
4134	902	537	CHINANET-BACKBONE
17908	599	536	Tata Communications
2706	546	528	Sino Software Research Centre
24560	660	479	Bharti Airtel Ltd.
4808	618	473	CNCGROUP IP network: China169
17676	527	466	Softbank BB Corp.
17974	488	463	PT TELEKOMUNIKASI INDONESIA
9443	495	413	Primus Telecommunications
7552	414	409	Vietel Corporation
10091	348	337	SCV Broadband Access Provider
4668	343	336	LG-EDS Systems Inc.

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

North America Aggregation Savings Summary

ASN	No of Nets	Savings	Description
209	2832	2218	Qwest
1785	1719	1562	PaeTec Communications, Inc.
20115	2032	1291	Charter Communications
4323	1617	1245	Time Warner Telecom
11492	1218	1206	Cable One
18566	1060	1050	Covad Communications
22773	995	938	Cox Communications, Inc.
5668	756	728	CenturyTel Internet Holdings,
6517	742	697	Yipes Communications, Inc.
6478	1192	691	AT&T Worldnet Services
2386	1581	685	AT&T Data Communications Serv
19262	940	680	Verizon Global Networks
3356	1115	672	Level 3 Communications, LLC
855	593	538	Canadian Research Network
33588	554	537	Bresnan Communications, LLC.
7029	551	462	Alltel Information Services,
7018	1438	438	AT&T WorldNet Services
11139	457	427	Cable & Wireless Dominica
19916	445	418	OLM LLC
10796	598	408	ServiceCo LLC - Road Runner

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

Latin America Aggregation Savings Summary

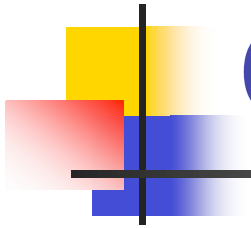
ASN	No of Nets	Savings	Description
8151	1473	1245	UniNet S.A. de C.V.
11830	677	668	Instituto Costarricense de El
10620	706	631	TVCABLE BOGOTA
22047	563	549	VTR PUNTO NET S.A.
7303	509	431	Telecom Argentina Stet-France
16814	426	416	NSS, S.A.
6471	438	397	ENTEL CHILE S.A.
28573	370	350	NET Servicos de Comunicao S.A
11172	408	334	Servicios Alestra S.A de C.V
7738	360	333	Telecomunicacoes da Bahia S.A
14117	338	329	Telefonica del Sur S.A.
10481	309	300	Prima S.A.
3816	330	264	Empresa Nacional de Telecomun
14522	244	234	SatNet S.A.
8167	257	223	TELESC - Telecomunicacoes de
14420	236	218	ANDINATEL S.A.
21826	248	217	INTERCABLE
19090	233	211	Canbras Net Ltda.
8163	207	199	METROTEL REDES S.A.
14259	279	197	GTD Internet S.A.

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

EU & Middle East Aggregation Savings Summary

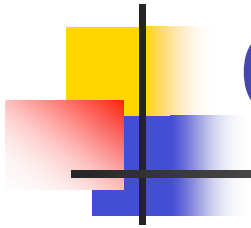
ASN	No of Nets	Savings	Description
8452	890	883	TEDATA
12479	393	387	Uni2 Autonomous System
29049	325	322	AzerSat LLC.
8866	333	311	Bulgarian Telecommunication C
35805	312	304	United Telecom of Georgia
8551	304	268	Bezeq International
9198	269	258	Kazakhtelecom Data Network Ad
9121	259	234	TTnet Autonomous System
3215	330	231	France Telecom Transpac
9155	212	203	QualityNet AS number
8877	205	198	BOL.BG Autonomous System
3352	237	194	Ibernet, Internet Access Netw
30890	387	188	SC Kappa Invexim SRL
3269	249	176	TELECOM ITALIA
12978	179	175	Dogan Iletisim Elektronik Ser
29314	183	171	Telewizja Kablowa Dami Sp. z
6830	190	147	UPC Distribution Services
12735	165	144	Netone Bilgi Ve Iletisim Hizm
3300	237	142	AUCS Communications Services
9051	157	132	INCONET Autonomous System

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