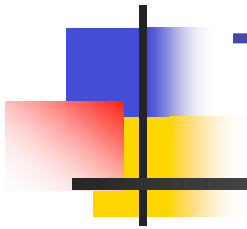
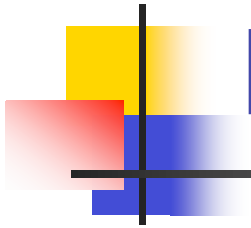


BGP Aggregation & The Deaggregation Report

Philip Smith

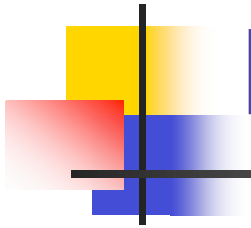
SANOG 11
10th-18th January 2008
Dhaka, Bangladesh

Route Aggregation Recommendations

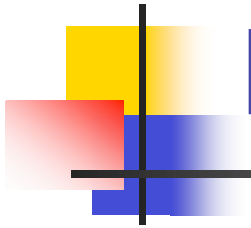


- LINX attempted aggregation policy for members
  - It failed even though most members voted for policy
- RIPE Routing Working Group work item from early 2006
  - Based on early LINX concept
  - Authored by Philip Smith, Mike Hughes (LINX) and Rob Evans (UKERNA)

Route Aggregation Recommendations

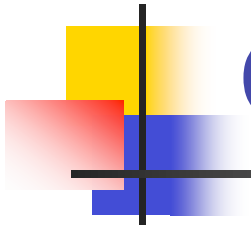


- RIPE Document — RIPE-399
  - http://www.ripe.net/ripe/docs/ripe-399.html
- Discusses:
  - History of aggregation
  - Causes of de-aggregation
  - Impacts on global routing system
  - Available Solutions
  - Recommendations for ISPs



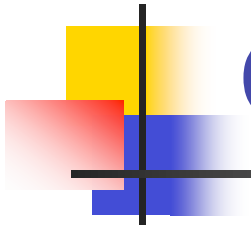
History:

- 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



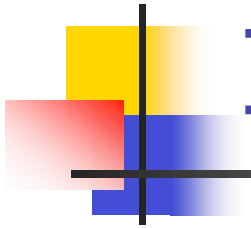
Deaggregation: Claimed causes (1):

- 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"



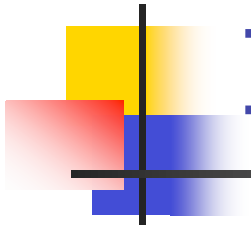
Deaggregation: Claimed causes (2):

- 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



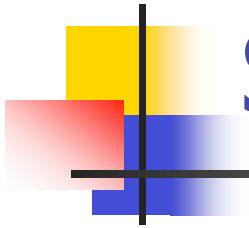
Impacts (1):

- 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



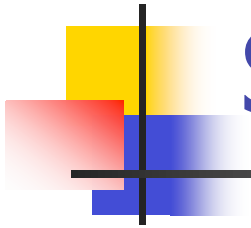
Impacts (2):

- 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



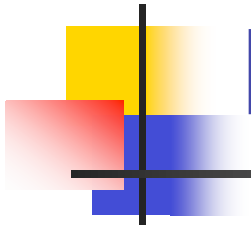
Solutions (1):

- CIDR Report
  - Global aggregation efforts
  - Running since 1994
- Routing Table Report
  - Per RIR region aggregation efforts
  - Running since 1999
- Filtering recommendations
  - Training, tutorials, Project Cymru,…
- “CIDR Police”



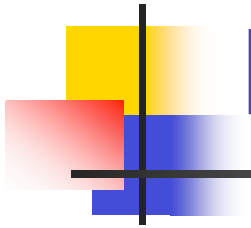
Solutions (2):

- 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



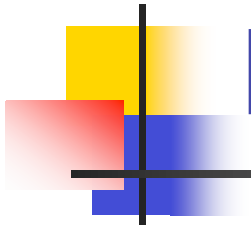
RIPE-399 Recommendations:

- 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
- (Oh, and all this applies to IPv6 too)



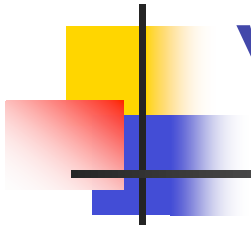
Looking at Deaggregation

- 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"



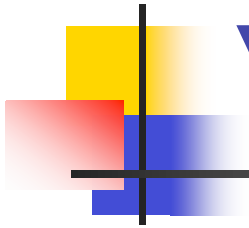
“Original Internet” — 2008/01

Total Prefixes

- Global BGP Table
  - 243k prefixes
- North America
  - 115k prefixes
- Europe & Middle East
  - 51k prefixes

Deaggregation Factor

- Global Average
  - 1.96
- North America
  - 1.76
- Europe & Middle East
  - 1.58



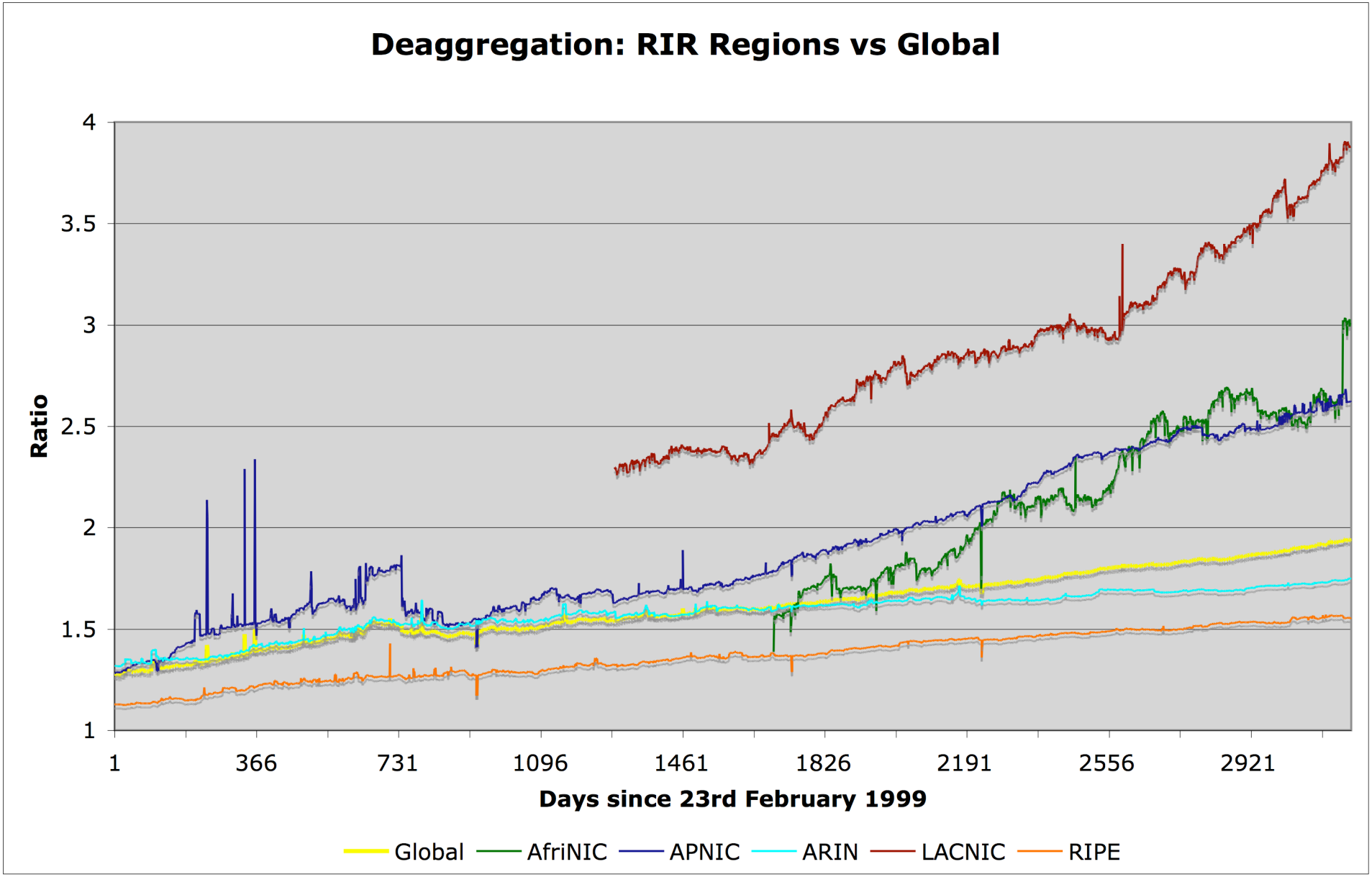
“Newer Internet” — 2008/01

Total Prefixes

- Global BGP Table
  - 243k prefixes
- Asia & Pacific
  - 56k prefixes
- Africa
  - 3k prefixes
- Latin America & Caribbean
  - 18k prefixes

Deaggregation Factor

- Global Average
  - 1.96
- Asia & Pacific
  - 2.66
- Africa
  - 3.05
- Latin America & Caribbean
  - 3.98


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

Africa Aggregation Savings Summary

ASN	No of Nets	Poss Savings	Description
24863	403	370	LINKdotNET AS number
20858	202	199	EgyNet
6713	145	134	Itissalat Al-MAGHRIB
33783	142	131	EEPAD TISP TELECOM & INTERNET
5536	124	114	Internet Egypt Network
33776	99	92	Starcomms Nigeria Limited
24835	88	82	RAYA Telecom - Egypt
15475	85	81	Nile Online
29571	77	70	Ci Telecom Autonomous system
20484	71	67	Yalla Online Autonomous Syste
23889	77	61	MAURITIUS TELECOM
3741	279	54	The Internet Solution
15706	57	53	Sudatel Internet Exchange Aut
21152	32	31	AS for the uplinks of Soficom
12455	32	29	Jambonet Autonomous system
2018	142	28	Tertiary Education Network
10798	27	26	Standard Bank of South Africa
8524	33	25	AUCEGYPT Autonomous System
15964	30	23	CAMEROON TELECOMMUNICATIONS N
33774	50	22	AS Number for Telecom Algeria

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

Asia & Pacific Aggregation Savings Summary

ASN	No of Nets	Poss Savings	Description
9498	1097	1036	BHARTI BT INTERNET LTD.
17488	960	873	Hathway IP Over Cable Interne
9583	1113	677	Sify Limited
9829	590	579	BSNL National Internet Backbo
4134	860	554	CHINANET-BACKBONE
18101	609	549	Reliance Infocom Ltd Internet
4780	591	548	Digital United Inc.
4668	521	511	LG-EDS Systems Inc.
4766	834	497	Korea Telecom (KIX)
4812	529	443	China Telecom (Shanghai)
17676	506	442	Softbank BB Corp.
4808	530	405	CNCGROUP IP network: China169
7545	497	404	TPG Internet Pty Ltd
17974	395	378	PT TELEKOMUNIKASI INDONESIA
9443	434	365	Primus Telecommunications
4538	358	320	China Education and Research
4802	471	319	Wantree Development
7552	298	294	Vietel Corporation
9929	331	281	China Netcom Corp.
9394	266	253	CHINA RAILWAY Internet(CRNET)

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

North America Aggregation Savings Summary

ASN	No of Nets	Poss Savings	Description
11492	1210	1194	Cable One
18566	1040	1030	Covad Communications
4323	1389	1026	Time Warner Telecom
6478	1096	964	AT&T Worldnet Services
22773	839	781	Cox Communications, Inc.
19262	868	717	Verizon Global Networks
5668	666	647	CenturyTel Internet Holdings,
6517	607	566	Yipes Communications, Inc.
15270	602	546	PaeTec.net -a division of Pae
2386	1354	531	AT&T Data Communications Serv
19916	563	521	OLM LLC
855	556	503	Canadian Research Network
7018	1467	473	AT&T WorldNet Services
6197	968	466	BellSouth Network Solutions,
6140	616	464	ImpSat
33588	474	442	Bresnan Communications, LLC.
3356	835	423	Level 3 Communications, LLC
7011	1004	414	Citizens Utilities
6389	439	404	bellsouth.net, inc.
20115	863	394	Charter Communications

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

Latin America Aggregation Savings Summary

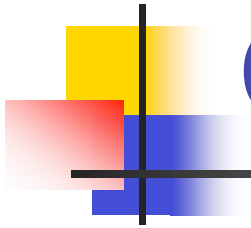
ASN	No of Nets	Poss Savings	Description
8151	1150	932	UniNet S.A. de C.V.
11830	545	536	Instituto Costarricense de El
16814	427	417	NSS, S.A.
7303	421	367	Telecom Argentina Stet-France
6471	379	342	ENTEL CHILE S.A.
14117	351	335	Telefonica del Sur S.A.
11172	396	334	Servicios Alestra S.A de C.V
10620	350	325	TVCABLE BOGOTA
22047	322	309	VTR PUNTO NET S.A.
10481	301	291	Prima S.A.
6147	255	234	Telefonica Del Peru
11556	236	232	Cable-Wireless Panama
7738	239	214	Telecomunicacoes da Bahia S.A
20299	236	207	NEWCOM AMERICAS
28573	228	205	NET Servicos de Comunicao S.A
23216	225	178	RAMtelecom Telecomunicaciones
19169	193	173	Telconet
14522	170	162	SatNet S.A.
8163	170	154	METROTEL REDES S.A.
10834	193	153	ADVANCE TELECOMUNICACIONES S.

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

EU & Middle East Aggregation Savings Summary

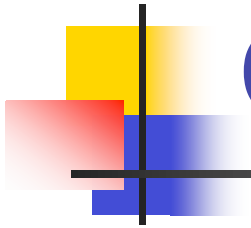
ASN	No of Nets	Poss Savings	Description
8452	287	280	TEDATA
5462	285	258	Telewest Broadband
8866	277	253	Bulgarian Telecommunication C
8551	277	236	Bezeq International
9155	215	205	QualityNet AS number
3352	240	200	Ibernet, Internet Access Netw
29357	199	195	WATANIYA TELECOM
12479	195	188	Uni2 Autonomous System
9121	209	184	TTnet Autonomous System
3215	270	178	France Telecom Transpac
9116	194	166	Goldenlines main autonomous s
3269	231	163	TELECOM ITALIA
6830	183	142	UPC Distribution Services
5486	154	136	Euronet Digital Communication
29049	130	128	AzerSat LLC.
5384	127	117	Emirates Internet
15471	171	116	SNR - Societatea Nationala de
3300	211	111	AUCS Communications Services
12883	99	95	Farlep-Internet ISP
31083	99	95	PowerNet.BG, Sofia, Bulgaria

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



Observations

- Huge gulf in operational good practices between “older” and “newer” Internet
  - Could threaten the Internet as we know it
- 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

- "Newer" Internet is growing rapidly
  - As is the deaggregation there
- RIPE-399 now exists
- Make it your BGP good practice document