What is going on with BGP

~ IETF 117

Perception of BGP standardization

- Completed long time ago, nothing to see here, end of story.

Perception of BGP standardization

RFC1105, Jun 1989 – BGPv1, the napkin FSM, short marker, link type

RFC1163, Jun 1990 – BGPv2, long marker, path attributes, origin control

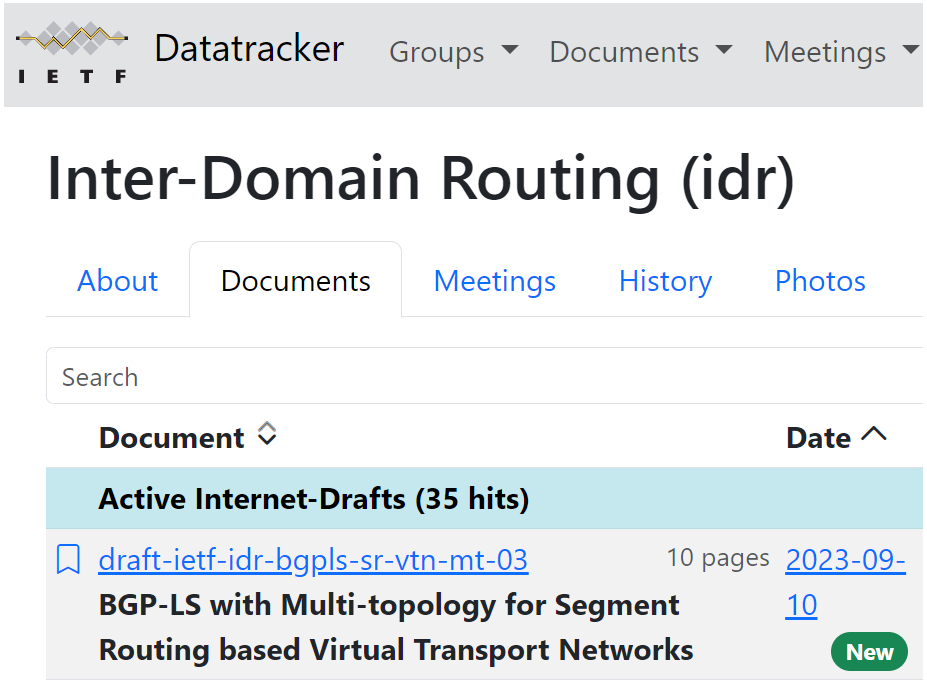
RFC1267, Oct 1991 – BGPv3, router identifier, third party nexthop

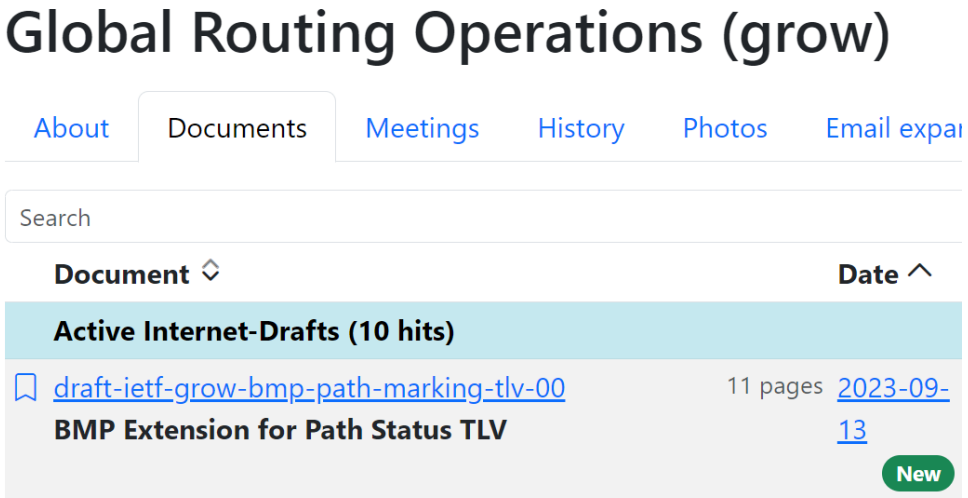
RFC1654, Jul 1994 – BGPv4, classless.

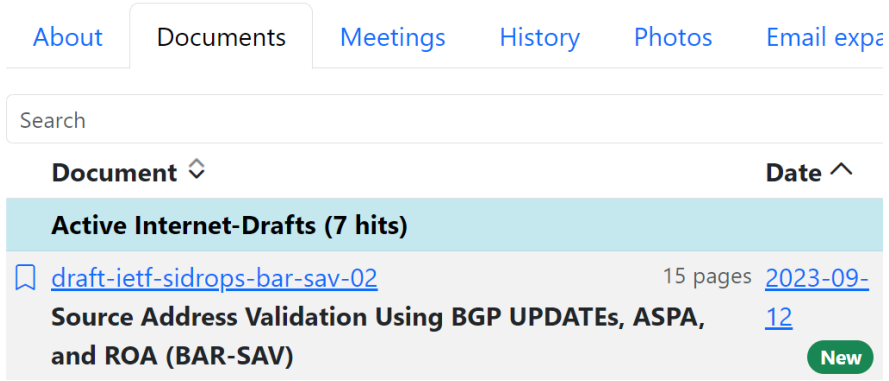
RFC1771, Mar 1995 – BGPv4, minor cleanup, aggregation.

RFC4271, Jan 2006 – BGPv4, major cleanup of 2002.

Perception of BGP standardization







BGP at the IETF

- idr – core protocol
- sidr -> sidrops – ROV, ASPA, BGPsec
- bess - EVPN
- grow – operations, BMP
- rtgwg – general topics on routing
- lsvr – experimental hybrid

Plus a set of other WGs that use BGP as a universal transport mechanism.

Standardizing BGP

- RFC published but no features?
- No RFC but features available?
- RFC says this but platforms implement that?
- Interoperability limits
- Featuritis
- Severe form of NIH syndrome
- Nonoverspecification

Deprecation of AS sets

draft-ietf-idr-deprecate-as-set-confed-set

- AS path contains segments
- Separate segment types for sequences and sets
- Sets are unordered
- Conflicts with origin validation
- Incompatible with path validation
- Deprecation of AS sets, not of aggregation itself!

Extended messages

RFC8654

- BGP PDU are not fragmentable
- Size of individual attributes may be large
- Total size of all attributes may be large
- Packing multiple prefixes is desirable
- 4096 -> 65535 octets
- Not applicable to OPEN messages
- Signalled as a capability
- NOTIFICATION may still be truncated

Extended optional parameters

RFC9072

- Capabilities are optional parameters
- Overall capability length is increasing
- A new optional parameter acting as a container for capabilities
- Backwards compatible
- ~4K octets available

Dynamic capabilities

draft-ietf-idr-dynamic-cap

- Capabilities are static for the duration of BGP session
- That may be too restrictive in some cases
- A new message type, used for two-way communication
- Negotiation of capabilities and their parameters
- Might get complex really fast

Send hold timer

draft-ietf-idr-bgp-sendholdtimer

- BGP hold timer is used for the receiving side only
- Complex cross-layer failures do happen
- Inability for the remote peer to process its incoming messages is a failure
- Send timer is a hold timer in reverse

Optimal route reflection

RFC9107

- Path selection decisions are local
- That may not be optimal from the perspective of RR client
- Specifically, IGP metrics most likely will be different
- A simple computational task, no protocol changes
- Has performance implications

Wide communities

draft-ietf-idr-wide-bgp-communities

- Haven’t we got enough of various communities already?
- Typing and parametrization
- Community container path attribute
- Wide community sub-objects
- Transitivity control
- Matching control
- Commuity:Source:Target:Parameters

Extended administrative shutdown

RFC9003

- Maintenance fix of RFC8302
- 128 -> 255 octets
- UTF-8

BFD cease

RFC9384

- Cease subcode to indicate session going down due to BFD

LLGR

draft-ietf-idr-long-lived-gr

- Long-lived GR
- Mostly a maintenance fix for initial oversight on value ranges
- 4095 seconds may seem long, but not long enough
- LLGR community to indicate that this prefix may be stale for a long time
- NO_LLGR community to avoid holding stale prefixes

Peer roles

RFC9234

- Adds relationship roles to BGP session peers
- A role defines and constrains what a peer can advertise and receive
- A capability and an additional path attribute
- All of this can and is implementable via policy only
- Policy still stays in place, roles just augment it

Well-known large communities

draft-heitz-idr-wklc

- Large communities are just a container – that is by design
- No propagation control and parameter classification
- Reuse current 4:4:4 structure as 2:10
- A controversial proposal indeed

Prepending considerations

draft-ietf-grow-as-path-prepending

- A writeup on well-known truths on the perils of prepending, specifically of excessive prepending
- Prepending is a powerful tool, and is often abused
- Prepending beyond a small count has no observable benefit but adds potential risk
- Overall a good example of a document that brings in operational value

Prefix limits

draft-sas-idr-maxprefix-inbound
draft-sas-idr-maxprefix-outbound

- Pre-policy and post-policy limiting
- Inbound and outbound limiting
- A simple yet valuable mechanism

BMP for local RIB

draft-ietf-grow-bmp-local-rib

- BMP originally provided for capturing of inbound and outbound RIB state
- Local RIB was out of scope, yet this may be beneficial
- Local RIB is treated as a separate peer
- No changes to BMP semantics as such

BMP TLVisation

draft-ietf-grow-bmp-tlv

- Initial design of BMP did not allow for decoration of route monitoring messages with additional attributes
- Deployments indicated for this to be needed
- Initial attempts resulted in fragmented point solutions that forecasted for conflicts and unjustified complexity
- Therefore a generalized TLV-based framework for all of BMP messages
- Ability for stateless parsing

ASPA

draft-ietf-sidrops-aspa-profile,
draft-ietf-sidrops-aspa-verification

- Origin validation is good but not enough
- Path validation is also required
- Relationship between ASes is known and can be used for building a list of adjacencies that can be verified later
- Reuse RPKI for storage and RTR for signaling

RTR v2

draft-ietf-sidrops-8210bis

- RTR extended for carrying information elements required for ASPA operation
- Some clarifications of dealing with multiple ROAs and prefix sorting
- Same operational model as for previous versions.

Alternative realities^wAFs

- EVPN and L3VPN interworking
- Flowspec maintenance
- BGP-LS
- Extensions for SR
- LSVR

Alternative transports

- Multisession
- BGP over QUIC

Open discussion