

Super-cluster with Gateways



- [Gateways](#)
 - [Gateway Connections](#)
 - [Interest Propagation](#)
 - [Interest-only Mode](#)
 - [Queue Subscriptions](#)
 - [Gateway Configuration](#)
-

Gateways

Gateways enable connecting one or more clusters together into a full mesh; they allow the formation of superclusters from smaller clusters. Cluster and Gateway protocols listen on different ports. Clustering is used for adjacent servers; gateways are for joining clusters together.

Gateway configuration is similar to clustering:

- gateways have a dedicated port where they listen for gateway requests
- gateways gossip gateway nodes and remote discovered gateways

Unlike clusters, gateways:

- have names, specifying the cluster they are in
- don't form a full mesh between gateway nodes, but form a full mesh between clusters instead
- are bound by uni-directional connections
- don't gossip gateway nodes to clients

Gateways exist to:

- reduce the number of connections required between servers
- optimize the interest graph propagation

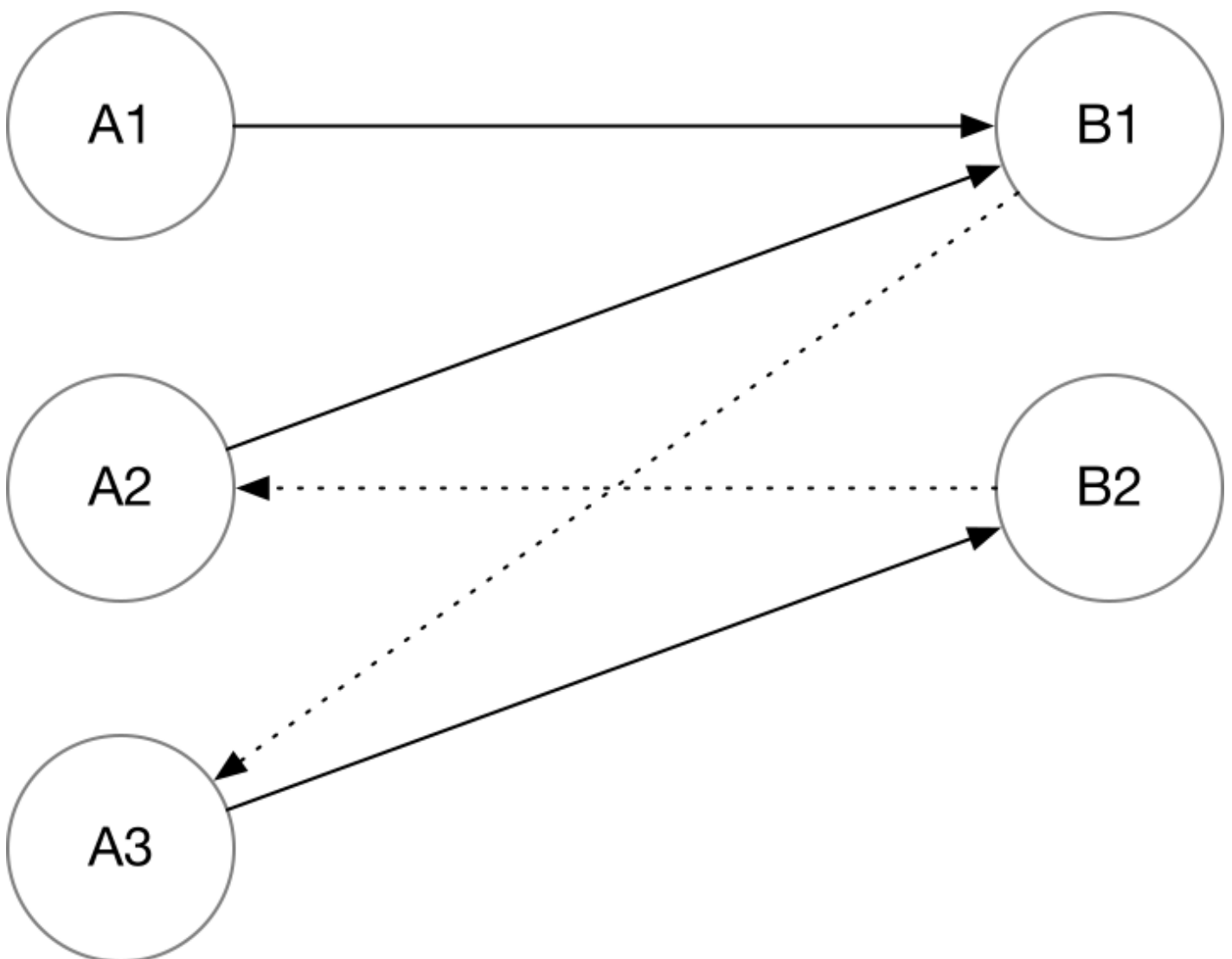
If gateways are to be used in a cluster, **all** servers of this cluster need to have a gateway configuration with the **same name**.

Furthermore, every gateway node needs to be able to **connect to any** other gateway node and vice versa. Everything else is considered a misconfiguration.

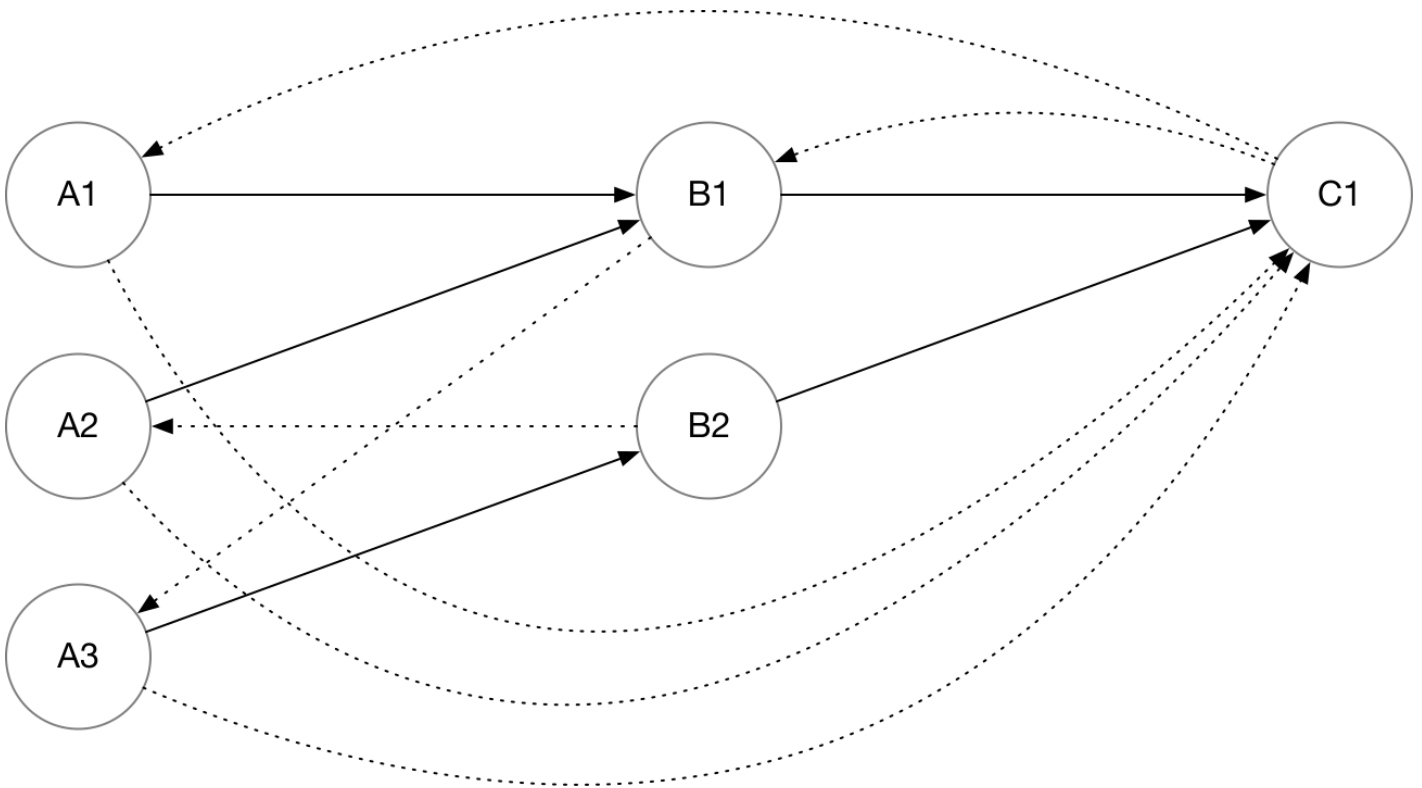
Gateway Connections

A nats-server in a gateway role will specify a port where it will accept gateway connections. If the configuration specifies other *external* `gateways`, the gateway will create one outbound gateway connection for each gateway in its configuration. It will also gossip other gateways it knows or discovers. Fewer *external* `gateways` mean less configuration. Yet, the ability to discover more gateways and gateway nodes depends on these servers running. This is similar to *seed server* in cluster. It is recommended to have all *seed server* of a cluster listed in the `gateways` section.

If the local cluster has three gateway nodes, this means there will be three outbound connections from the local cluster to each external gateway cluster.



In the example above cluster A has configured gateway connections for B (solid lines). B has discovered gateway connections to A (dotted lines). Note that the number of outgoing connections always matches the number of gateways with the same name.



“ In this second example, again configured connections are shown with solid lines and discovered gateway connections are shown using dotted lines. Gateways A and C were both discovered via gossiping; B discovered A and A discovered C.

A key point in the description above is that each node in the cluster will make a connection to a single node in every remote cluster — a difference from the clustering protocol, where every node is directly connected to all other nodes.

For those mathematically inclined, cluster connections are $\frac{N(N-1)}{2}$ where N is the number of nodes in the cluster. On gateway configurations, outbound connections are the summation of $N_i(M-1)$ where N_i is the number of nodes in a gateway i , and M is the total number of gateways. Inbound connections are the summation of $U - N_i$ where U is the sum of all gateway nodes in all gateways, and N is the number of nodes in a gateway i . It works out that both inbound and outbound connection counts are the same.

The number of connections required to join clusters using clustering vs. gateways is apparent very quickly. For 3 clusters, with N nodes:

Nodes per Cluster	Full Mesh Conns	Gateway Conns
1	3	6
2	15	12
3	36	18
4	66	24
5	105	30
30	4005	180

A cluster section is not needed for gateways, they work with single server as well. Yet, they start to be useful when participating cluster consist of more than one server and they reduce the number of connections.

Interest Propagation

Messages from clients directly connected to a gateway node will be sent along outgoing gateway connections according to the following two interest propagation mechanisms:

- Interest-only Mode
- Queue Subscriptions

Local interest permitting, the receiving gateway node sends the messages directly to its subscribing clients as well as the server within the cluster.

Interest-only Mode

Gateway *A* sends messages only for subjects that gateway *B* has explicitly expressed interest in. As subscriptions come and go on *B*, *B* will update its subject interest with *A*.

Queue Subscriptions

When a queue subscriber creates a new subscription, the gateway propagates the subscription interest to other gateways. The subscription interest is only propagated *once* per *Account* and subject. When the last queue subscriber is gone, the cluster interest is removed.

Queue subscriptions work on *Interest-only Mode* to honor NATS' queue semantics across the *Super Cluster*. For each queue group, a message is only delivered to a single queue subscriber. If the same queue group exists in multiple clusters, the server will pick one member from the queue group in its cluster, only sending to a different cluster if there is no interest in its cluster. In other words, a server will always try to serve local queue subscribers first and only failover when a local

queue subscriber is not found. The server will pick the cluster with the lowest RTT.

Gateway Configuration

The [Gateway Configuration](#) document describes all the options available to gateways.

[] [] #4	
[] cocobook-sync [] []	2026-10-28 11:13:48 UTC
[] cocobook-sync [] []	2026-10-29 03:58:10 UTC