Matrix

- State resolution attacks

State resolution attacks

These are some notes on various different kinds of attacks that might be attempted on state resolution algorithms, such as the one in Matrix. Different kinds of state resolution algorithms are vulnerable to different kinds of attacks; a reliable state algorithm should be vulnerable to none of them.

> These notes are not complete. More details, graphs, etc. will be added at some later time.

Frontrunning attack

Detect an event that bans or demotes the user, then quickly craft a fake branch full of malicious events (eg. banning other users), but do not submit those events to any other homeserver yet, and then craft an event that parents both the fake branch and the event prior to the detected ban/demote, claiming that the fake branch came earlier and thereby bypassing the ban. Requires a malicious homeserver.

Dead horse attack

Attach crafted event to recent parent and ancient parent, to try and pull in ancient state and confuse the current state; eg. an event from back when a user wasn't banned yet, to try and get the membership state to revert to 'joined' by pulling it into current state. Named this because it involves "beating a dead horse".

Piggybacking attack

A low-powerlevel user places an event in a DAG branch that a high-powerlevel user has also attempted to change state in, as the high-powerlevel state change might cause their branch to become prioritized (ie. sorted in front) in state resolution.

Fir tree attack

Resource exhaustion attack; deliberately constantly creating side branches to trigger state resolution processes. Named after the shape of half a fir tree that it generates in the graph.

Huge graph attack

Resource exhaustion attack; attach crafted event to a wide range of other parent events throughout the history of the room, to pull as many sections of the event graph into state resolution as possible

Mirror attack

Takes advantage of non-deterministic state resolution algorithms to create a split-brain situation that breaks the room, by creating a fake branch containing the exact inverse operations of the real branch, and then resolving the two together; as there is no canonically 'correct' answer under these circumstances, the goal of the attack is to make different servers come to different conclusions.