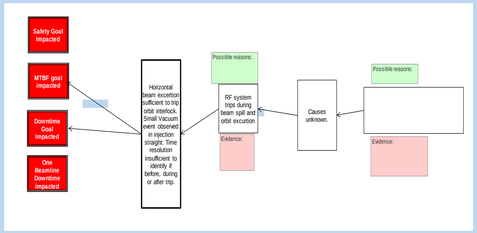
Root Cause Analysis Experience, Highs and Lows.

Christopher Bailey, Diamond Light Source

Overall since starting to use RCFA on all trips we have seen our annual MTBF increase from around 24 hrs to above 100 hrs.
However we have mixed experience with reactions to requiring a root cause failure analysis is performed on all major systems failures on Diamond. Ideally we get clear identification of the causes and a practical set of actions to address them, but this is not always the result. The less successful results we can get are either :- An over simple evaluation of problem, or a failure to focus on key issues resulting in a very long list of speculative actions, or an assumption that the problem has already been looked at as far as possible.
Which leads to the question how can we improve the way we run the process to continue to get a useful analysis more often?

Suggestions:-
Ensure all diagnostics are working.
Suggest new monitoring methods
But need to have good idea where

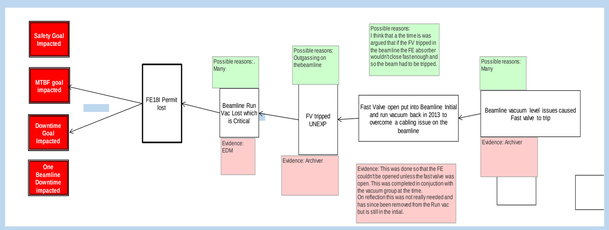
Unable to find enough information about the cause of the trip!



Cause #	Action Item to rectify Cause(s)	Owner	Due Date	Status	Completed	Comments
C1	None until causes can be identified!					

Are we confident there isn't another route to this problem?

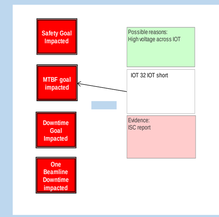
This one gave a clear action, with the single action sufficient to stop recurrence.



Cause #	Action Item to rectify Cause(s)	Owner	Due Date	Status	Completed	Comments
C1	Remove FE open signal from Run Vac	SL	Immediately	Complete	29-Sep	This is the only beamline with a Fast valve close enough to the FE absorber to cause this issue.

We need to encourage thinking further around the problem; Are possibilities avoided because they look too large?

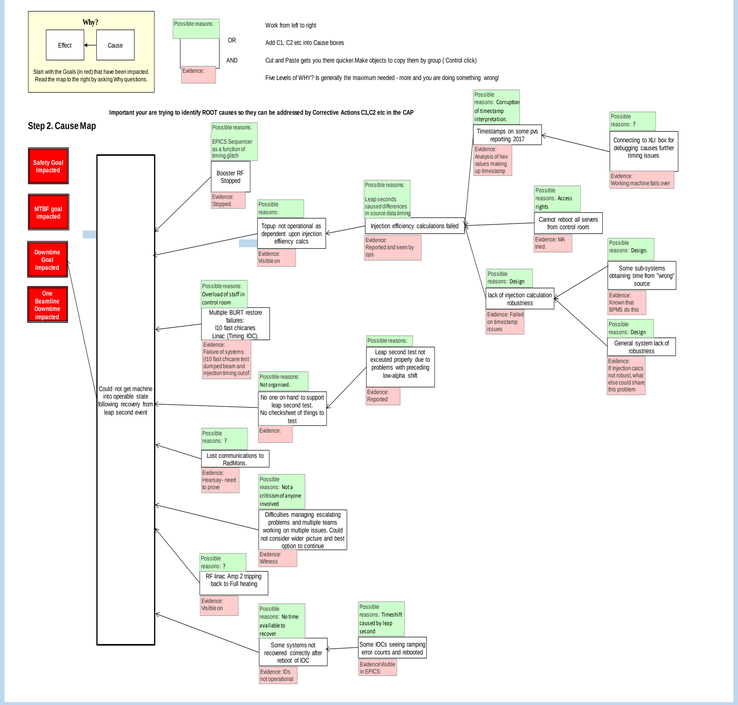
Given up, we've seen this too often before and have run out of ideas



Cause #	Action Item to rectify Cause(s)	Owner	Due Date	Status	Completed	Comments
A1	OT 3/2 HV connectors inspected	PM	11/07/2017	Not started		

Possibly we should have looked parallel lines separately.

Detailed Analysis but too much brought forward as actions



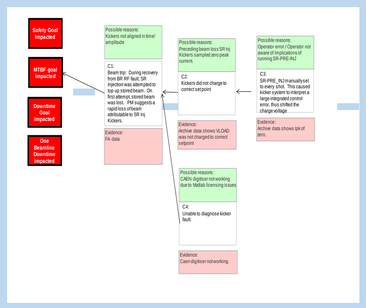
The aim using root cause fault analysis is to identify actions, which reduce the rate of failure or improve the recovery time.
These need to be possible to implement in a reasonable timescale and cost.
They don't necessarily need to prevent the problem just to stop it escalating by breaking the chain of events. When we conclude we have insufficient information actions to improve data capture are useful.

The Analysis can be poor for the following reasons:-
The fault is treated as the same as a previous one and not really analysed.
There is insufficient information to work out what happened and so cannot suggest any actions.
The Author can analyse the situation effectively proposing a redesign of the entire system rather than giving any actions that can be achieved promptly.

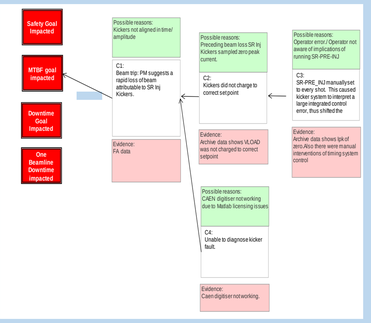
The Remaining questions are what can we do to encourage the authors of the analyses to find appropriate actions?
Also do we need to admit that sometimes it might be impossible to get to the root cause of the problem?
Even if we had better instrumentation and logging.
How do we keep that requirement for these analyses positive and avoid the feelings that it is generating a blame culture?

A Second similar but different fault may give us some additional information about the mechanisms involved?

Two very Similar faults that have some causes and effects in common. The actions described do result in a more robust system, but the 2nd is nearly a repeat of the first and does not look at the differences in the situation.



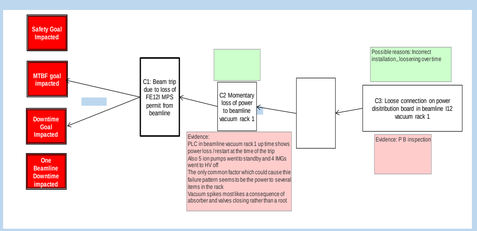
Cause #	Action Item to rectify Cause(s)	Owner	Due Date	Status	Completed	Comments
A1	Meeting with A.Rose & P.Hamadyk to discuss possible causes of fault	C.A	19/01/2016	Complete	21/01/2016	
A2	Setup Vcharge interlock on top-up application	C.A, N.B	03/02/2015	Complete		
A3	Document cause and effect to circulate to relevant parties	P.H	26/01/2016	Complete	27/01/2016	
A4	Interlock SIK timing events with Make all Setpoints zero control (which connects to the top up control).	P.H	01/04/2016	Complete		Tested
A5	Review all SR injection timing events.	A.R/ P.H	01/04/2016	Complete	01/02/2016	Both EVRs had zero receiver error counts
A6	Implement CAEN digitiser logging system in something else other than Matlab, as this ties up a Matlab license.	A.R	01/04/2016	Complete		Soft IOC monitoring kicker waveforms. Data saved on beam dump during injection or current loss.



Cause #	Action Item to rectify Cause(s)	Owner	Due Date	Status	Completed	Comments
A1	Meeting with A.Rose & P.Hamadyk to discuss possible causes of fault	C.A	19-Jan	Complete	21-Jan	All these actions are duplicate of CAP - 6676
A2	Setup Vcharge interlock on top-up application	C.A, N.B	03/02/2015	Complete		All these actions are duplicate of CAP - 6676 Released a new TopUp on 16/02 which included the changes required to apply tight tolerances to the SR Kicker magnets. With this change TopUp is now able to detect when the kickers have not ramped correctly, e.g. after firing while ramped down, and drop to a Failed state rather than kick the beam. This releases also resolves the front end dependency issue (FaultID 6703), but I'm not sure that was ever on a CAP.
A3	Document cause and effect to circulate to relevant parties	P.H	29/01/2016	Complete	27/01/2016	All these actions are duplicate of CAP - 6676
A4	Interlock SIK timing events with Make all Setpoints zero control (which connects to the top up control).	P.H		Complete		All these actions are duplicate of CAP - 6676
A5	Review all SR injection timing events.	A.R/ P.H	Apr-16	Complete		All these actions are duplicate of CAP - 6676
A6	Implement CAEN digitiser logging system in something else other than Matlab, as this ties up a Matlab license.	A.R/ P.H	Apr-16	Complete		All these actions are duplicate of CAP - 6676
A7	Education of operators with regards to function of the timing interface and impact on the control systems.	V. K	Apr-16	Complete		
A8	Investigate operation of the power controller sample/hold mechanism	C. A	Apr-16	Superceeded		Should the sample and hold mechanism have operated given that the initial conditions were incorrect?

Procedures at times long before the event might also need to be looked at.

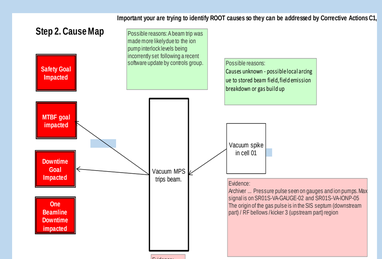
Power feeds dual routed or using UPS can still be insufficient if there are poor connections.



Cause #	Action Item to rectify Cause(s)	Owner	Due Date	Status	Completed	Comments
C3	Inspect and tighten connections on power distribution board in BL I12 vacuum rack 1	PB	24/06/2015	Complete	24/06/2015	
A1	Schedule electrical inspection of beamline racks and power connections	??	??	TAG		Working group aleready considering this - Andy Reid?
A2	Review status of duplex power supplies upgrade on beamlines. Not a direct cause here but another possible avoidable beam trip mechanism	AD - Beamlines.	End of 2015	TAG		New PSUs available - beamlines need to install them. Sent chase-up to Andy Dent on 9th September 2015:One of the outcomes of the investigation to the BL12I beam trip was to free-issue duplex power supplies to all beamlines to upgrade their vacuum #1 racks. Please see link below for details. "S:\Technical\Operations\Fault Log DB\Fault report Corrective action plans\2015 Run 3\6435 Beam trip - vacuum event on BL12I.xlsx" Except for I12 & I13 the uptake of this has been minimal. Please can you encourage beamlines to progress this. Simon Ley can advise how to go about installation. This work should really be confined to shutdowns, as loss of power to a vacuum IOC can cause top-up to stop due to loss of EPICS comms to the radiation monitors.

The problem isn't always the problem!

Acceptable solution had previously been implemented but lost on update.



Cause #	Action Item to rectify Cause(s)	Owner	Due Date	Status	Completed	Comments
C1	Review and increase local vacuum interlock levels	H Shiers	18/07/2017	Complete	18/07/2017	Ion pump interlocks wrer effecigve defeated follinwg a software update by controls group. Levels set manually as in elog entry H Shiers 18/7/17
C2	Investigate further - try to establish root cause	M Cox	21/07/2017	Complete	21/07/2017	Looking at IMG and ion pump archive data, the pressure spike is in the SIS septum downstream end / downstream bellows / kicker 3 vessel upstream end region. Event is not synchronised with injection. Pressure spikes of different sizes in this region do occur from time to time sporadically plus small spikes every time beam is dumped. Reasons unknown. They do not usually trip the beam. If the ion pump interlock levels had been set correctly then there would probably have been no beam trip this time also and operations could have continued uninterrupted.
C3	Correct error in ion pump control software so that interlocks levels will be set from database as before	A Rose	31/07/2017	Complete	31/08/2017	

Step 4 Corrective Action Plan - Action Root Cause Analysis final stage 06/10/2017 14:06

Fault No = 6447 and 6445 **Date of trip** 01-Jul-15 08:30 **Downtime** 3.00
Fault title Topup and bunch alignment Controls / Fault Status Open
Summary comments Leave until assessed at AOSPUM

Cause #	Action Item to rectify Cause(s)	Owner	Due Date	Status	Completed	Comments
C1	Booster RF Stopped	PH		Complete		
C2	Topup not operational as dependent upon injection effiency calcs	MA				See C14
C3	Multiple BURT restore failures: I10 fast chicanes Linac (Timing IOC).	MH/AR		Complete		
C4	No one on hand to support leap second test. No checksheet of things to test	MH/AR		Complete		
C5	Leap second test not exceuted properly due to problems with preceding low-alpha shift	N/A				
C6	Lost communications to RadMons.	PH		Decided Against		
C7	escalating problems and multiple teams working on multiple issues. Could not consider wider picture and best option to continue	CB/Ops		TAG		Action against CB
C8	RF linac Amp 2 tripping back to Full heating	PH		Complete		
C9	Some systems not recovered correctly after reboot of IOC	AR/CB/Ops		TAG		
C10	Injection efficiency calculations failed	MA		No Action.		
C11	Some IOCs seeing ramping error counts and rebooted	AR/PH		Complete		
C12	Timestamps on some pvs reporting 2017	PH		Complete		
C13	Cannot reboot all servers from control room	MH		TAG		
C14	lack of injection calculation robustness	MA		No Action.		
C15	Connecting to XLI box for debugging causes further timing issues	PH		No Action.		
C16	Some sub-systems obtaining data from "wrong" source	PH/MA/MH		In Progress		
C17	General system lack of robustness	AR/MH		TAG		
C18						
A1	Investigate why Booster RF Stopped	PH	01-Aug	Complete		Confirmed due to leap second.
A2						
A3	Improve understanding of BURT and improve robustness and visibility	MH/AR	01-Sep	Complete		Closed. Mauro Faria has released a first version of a BURT status screen. This will change and grow. It is also an action on Ops to make use of it.
A4	Ensure coverage in control room	MH/AR		Complete		Mark and Austen are aware that there should have been better coverage in the control room from the controls group and will work to ensure that we have a better response to this.
A5						
A6	Establish whether there was lost communications to RadMons.	PH	01-Aug	Decided Against		Not sure whether this is true
A7	Review management in control room during major system failures	CB/Ops	01-Sep	TAG		Recovery strategy: There was a lot going on with people working on multiple different systems. Not clear how to proceed or priority. A review might help?
A8	Establish why RF linac Amp 2 tripping back to Full heating	PH	01-Aug	Complete		CC fixed this and it was due to the leap second.
A9	Write document to outline recovery steps for sub-systems	AR/CB/Ops	01-Sep	TAG		
A10	Full analysis of why injection efficiency calculations failed	MA	01-Aug	No Action.		Exactly why?
A11	Establish importance of error counts to IOCs	AR/PH	01-Aug	Complete		This is not an issue for almost everyone - do not restart IOCs.
A12	How did some timestamps on some pvs report 2017	PH	01-Aug	Complete		This has now been fixed in latest mrfTiming driver.
A13	Why can we not reboot all servers from control room	MH	01-Aug	TAG		More knowledge and who to call. Meeting organised with interested parties on 2nd December
A14	Improve injection calculation robustness	MA	01-Aug	No Action.		Possibly have a high integrity link
A15	Establish why connecting to XLI box for debugging causes further timing issues	PH	01-Aug	No Action.		Connecting to the XLI box caused another drop-out of timing. Understand relationship with EVG. Not reproducable.
A16	Determine which sub-systems are obtaining time from "wrong" source, document and advise on changes	PH/MA/MH	01-Sep	In Progress		Re-think overall architecture - long term. PH Reports: Only diagnostics rely on time stamps and, as discussed in the leap second meeting, I have fixed the problem with event time stamps in the future.
A17	Look for further system lack of robustness	AR/MH	01-Oct	TAG		Improve robustness across all critical control systems. In progress, but is not a one-time fix. Perhaps moved out of the CAP?
A18	Determine actions for next leap second	MH	01-Oct	TAG		Consider differerent approaches for dealing with leap seconds to improve overall robustness. Meeting organised for January 2016.
A19	Investigate leap second simulation	PH	01-Dec	Decided Against		PH Reports: Still investigating. I am rewriting the time stamp generator software and have found the software and development board.
A20	Find notification of leap seconds from central agency	PH	01-Dec	Complete		PH Reports: No more leap seconds this year. I didn't find a way to register to receive notifications on the main web sites but I am looking at a way to automate reading the notification.

diamond

ARW 2017
Accelerator Reliability Workshop