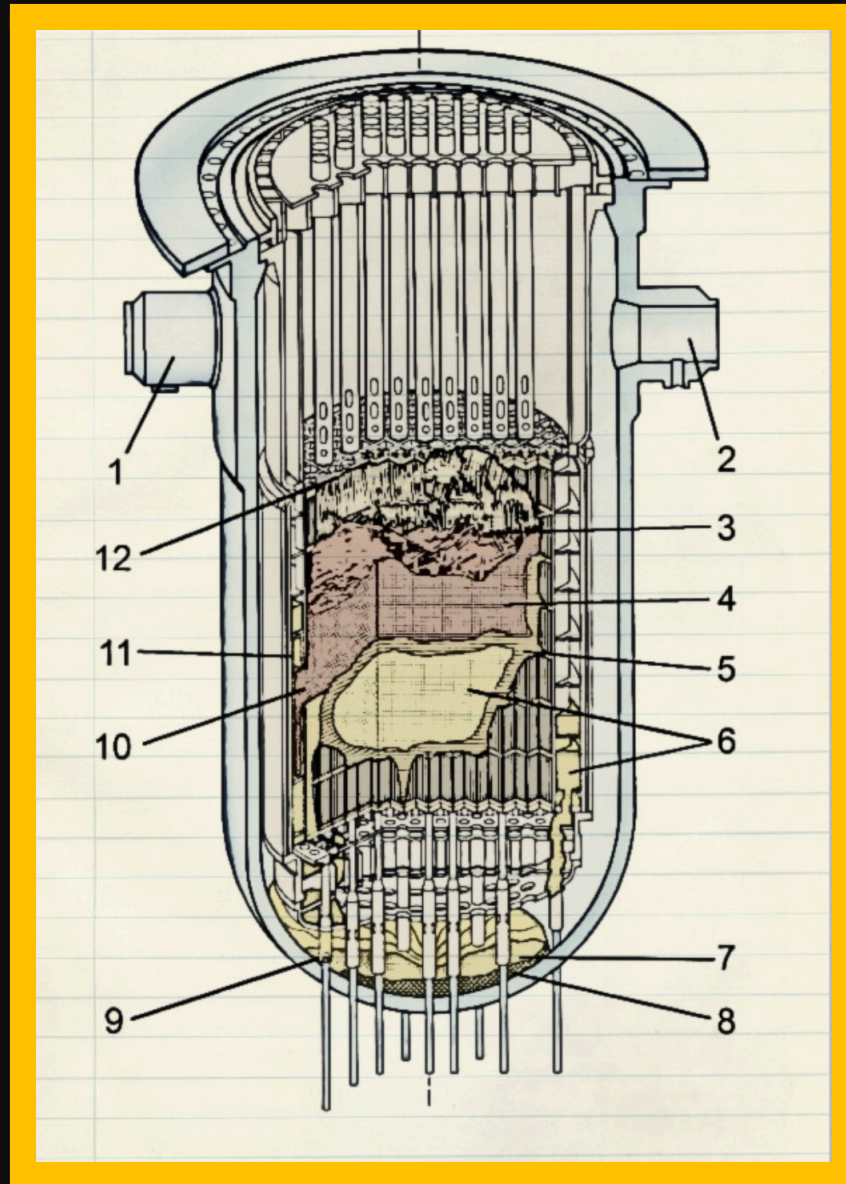


A Practical Analysis of Cyber-Physical Attacks Against Nuclear Reactors.



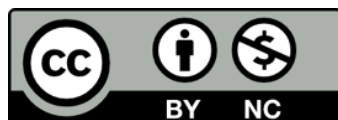
Contents

Author's Note	3
Scope	4
Executive Summary.....	5
1. Introduction.....	6
1.1 Nuclear Fission	7
1.2 When physics and engineering meet: Nuclear Reactors	8
1.3 Physics of Pressurized Water Reactors	9
1.3.1 - Neutron interactions with matter	9
1.3.2 - Microscopic Cross Sections	12
1.3.3 - From one to many	14
1.3.4 - Time matters	17
1.3.5 - Slow down, please.	18
1.3.6 - Nuclear fuel	20
1.3.7 - Reactivity Control	23
1.4 Nuclear Power Plants	38
1.4.1 - Safety	42
1.4.2 - Instrumentation and Control	45
2. Actors and motivations	49
2.1 Stuxnet.....	50
2.2 Trisis	52
2.3 Targets.....	54
3. Teleperm XS.....	55
3.1 Public availability of Teleperm XS	56
3.2 An overview of TXS	62
3.3 Exploring the attack surface	64
3.3.1 - Tianwan NPP and Russia's AES-2006	65
3.3.1.1 - Reactor Protection System / ESFAS	67
3.3.1.2 - I&C Service Center.....	69
3.3.2 - Oconee NPP	72
3.3.2.1 - The case of the "Staff Position 10"	75
3.3.3 - Research Reactors	86
3.4 Trisis-like attacks against the TXS platform.....	89

3.4.1 - The 'When'	92
3.4.2 - The 'Why'	93
3.4.3 - The 'How'	96
4. Cyber-Physical Attacks	97
4.1 Introduction	97
4.2 Preparation	101
4.3 Implementation	102
4.4 Sample cyber-physical attack: SLOCA via Pressurizer's PSRVs	109
4.4.1 - NeutronMode-IV NPP	109
4.4.2 - "Cyber Three Mile Island"	116
4.4.3 - No pressure.....	117
4.4.4 - Initiating event.....	119
4.4.5 - A matter of priorities	124
4.4.6 - Escalating a SLOCA into a severe accident	127
4.4.7 - Simulating the cyber-physical attack.....	130
5. Conclusions.....	134
About the author	135
Appendix A	136

October 1st, 2024.

This paper is released under the following license:



<https://creativecommons.org/licenses/by-nc/4.0/>

Author's Note

Currently humanity is facing significant challenges. One of the most pressing issues, as a direct result of anthropogenic activities, is the unprecedented climate conditions our planet is experiencing. We are now irredeemably ushering in the era of consequences, after collectively dismissing decades of clear, loud, and substantiated warnings coming from an overwhelming scientific consensus.

One of the lessons we are now painfully learning is that the control of atomic energy, one of the greatest achievements in scientific progress, was phased out due mainly to psychological fears and economic incentives, rather than solid scientific reasons. It would, however, be equally naive to dismiss, as a contributing factor to this situation, the perceived, and sometimes well-deserved, atmosphere of obscurity and secrecy that surrounds part of the nuclear industry.

The fact is that long-term damage, derived from the use of fossil fuels, was globally deemed an acceptable risk, as it was seen as a problem solely impacting future generations. In the meantime, nuclear energy was subjected to a gradual discredit. In certain countries, such as Germany, this resulted in the permanent shutdown of all of its nuclear power plants.

New advancements in fusion and fission technologies, such as Small Modular Reactors (SMRs), in addition to the undeniable reality we are living, seem to be paving the way for a 'nuclear renaissance'. I find this perspective encouraging as I honestly consider nuclear energy an essential asset for a sustainable future. However, for this endeavor to be successful, multiple challenges will need to be addressed. Cybersecurity is one of them, but also transparency and education.

Unfortunately, on top of these considerations, the current unstable geopolitical situation has once again, after many years, brought back the everlasting threat of a nuclear conflict.

Any cyber-attack that targets the safety systems of a Nuclear Power Plant (NPP) should be considered an extreme case, even for confrontations involving nation-state actors. Due to its potential consequences, it is only imaginable under the mandate of a strategic effort for destruction in preparation for, or during, an armed conflict.

In this context, I am presenting this research paper, which advocates for education as a vital tool for increasing public understanding of nuclear energy, with the intention to contribute to a better awareness of how cyber-physical attacks may impact nuclear facilities, driven from the perspective of hypothetical, but realistic, state-sponsored operations.

Ruben Santamarta.

Scope

This research has some inherent limitations due to the sensitive nature of the subjects covered, in addition to the restricted commercialization of software, hardware and in general terms, all the equipment used within the nuclear industry. In order to overcome these constraints, a considerable effort has been made to verify, by using different methods, all the technical details herein described.

The main objective of this publication is to provide a comprehensive technical analysis of hypothetical cyber-physical attacks targeting the safety systems of NPPs, for example, the Reactor Protection System (RPS) and Engineered Safety Features Actuation System (ESFAS). One of the novel aspects of this research is that it is based on the analysis of a digital safety Instrumentation and Control platform, Teleperm XS, that is currently deployed in multiple NPPs across Europe, USA, Russia and China.

This paper is structured to facilitate reading and understanding by a variety of readers, although a minimal technical background is assumed.

- The “*Introduction*” describes the nuclear engineering and nuclear physics concepts behind nuclear fission, Pressurized Water Reactors (PWRs) and NPPs, which are required to follow the subsequent cyber-physical attack scenarios. Prior knowledge of nuclear physics or reactor engineering is not assumed, making it accessible to those without a formal background in these fields.
- “*Actors and motivations*” describes the background of certain real-world operations involving cyber-physical attacks and nuclear facilities.
- “*Teleperm XS*” introduces the commercial Instrumentation and Control (I&C) platform, including a detailed description of the hardware, software architecture, attack surface, and eventually those characteristics that could potentially be leveraged by malicious actors.
- “*Cyber-Physical Attacks*” details an approach to analyzing the design of specific nuclear reactors in order to characterize a series of feasible cyber-physical attacks against their safety systems (e.g. RPS, ESFAS), according to the level of damage sought by the attackers.

Executive Summary

This research paper aims to provide a comprehensive technical analysis of hypothetical cyber-physical attacks targeting the safety systems of nuclear reactors (PWRs), such as the Reactor Protection System (RPS) and the Engineered Safety Features Actuation System (ESFAS).

To make this analysis accessible to readers with varying levels of technical expertise, the introduction provides a foundational understanding of the key concepts in nuclear engineering and nuclear physics that are essential for comprehending the subsequent attack scenarios.

The research focuses on Teleperm XS (TXS), a digital safety instrumentation and control platform used in multiple nuclear power plants (NPPs) in Europe, USA, China and Russia. It specifically explores the feasibility of Trisis-like attacks against TXS.

The attack scenarios analyzed in this paper are characterized by three main traits: they are remote, malware-based, and assume the attackers have already gained initial access to the non-safety network. A multi-step approach to executing such a sophisticated attack is elaborated on in this paper, which is outlined as follows:

1. Target Identification & Reconnaissance: Attackers must first identify vulnerable targets within the I&C system, gathering extensive intelligence on the reactor's design, operational procedures, and safety protocols.
2. Compromising the Service Unit: The Service Unit, with its privileged access to safety-critical components, emerges as a primary target. Attackers aim to gain control of this unit to send malicious commands and bypass security measures.
3. Exploiting System Weaknesses: Specific design issues are highlighted, such as the "Staff Position 10" requirement, which exposes a lack of hardware-enforced logic for changing operating modes in safety controllers. This feature, as demonstrated by the Trisis malware, could enable attackers to circumvent physical security measures like key switches.
4. Manipulating Safety Systems: With control over key components and by leveraging their priority levels, attackers can proceed to manipulate reactor parameters, inhibit crucial safety features in the Reactor Protection System (RPS) or Engineered Safety Features Actuation System (ESFAS), and interfere with operator actions. This, coupled with potential manipulation of information displayed to operators, could lead to misinterpretations, delayed responses, and a dangerous escalation of events.
5. Triggering a Physical Incident: Different attack scenarios are explored, such as triggering a Small Loss of Coolant Accident (SLOCA) by manipulating the pressurizer's Pressure Safety Relief Valves (PSRVs). This specific scenario leverages the complex interplay between digital commands, inherent physical processes, and operator responses, illustrating how a well-orchestrated cyberattack could lead to a core uncovering and subsequent fuel damage.

1.Introduction

A common job interview question for cybersecurity roles used to be something like:

“Could you please elaborate what happens when you type a url in your browser and hit enter?”.

Technically, this question can only be properly answered by writing dozens of books, it's just not possible to explain everything in detail without spending a prohibitive amount of time. However, the expected answer should condense the most important concepts and ideas of the underlying technology, and components involved during that process.

What if the question was something like this:

“Could you please elaborate what happens when a neutron induces a fission reaction inside a nuclear reactor?”

This is precisely what this section is all about.

Everything around nuclear physics may seem daunting at first, but if you are reading this research, there are two options:

1. You already know the answer, so you can skip this section.
2. Regardless of your career path or formal education, you already know the basis for understanding why nuclear reactors work, even if you were not aware of it. Please, let me explain myself.

At some point in our lives, we have all come across Einstein's famous equation $E = mc^2$, which settled the mass-energy equivalence.

While you're reading this, you don't expect either your mobile phone or your computer to spontaneously disintegrate. You are likely also pretty sure that your coffee mug will not fly across the room all of a sudden. Fortunately for us, in the physical world, despite all the complex interactions happening at different levels, there is always some predictability in them. A well-known statement that will likely sound familiar to everyone is '*energy is neither created nor destroyed*', which is another way to define this kind of predictability over time, best known as '*conservation*' in physics.

Among other things, a nuclear reactor is a mesmerizing, practical demonstration of these two big concepts: the mass-energy equivalence and its conservation.

Let's see why.

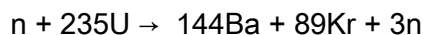
1.1 Nuclear Fission

In 1938 a group of German chemists discovered that by bombarding uranium –a heavy element– with neutrons, they were obtaining Barium, which was much lighter. That situation was so unexpected that they initially almost refused to believe it.

Some time later, Lise Meitner, an Austrian physicist exiled in Sweden due to persecution by the Nazis, provided a plausible explanation for the observed phenomenon. She elaborated the analysis in an historic letter to Nature¹, where the liquid-drop model and the mass-energy equivalence were pivotal elements to sustain this theory, as we can see in the following paragraph:

It seems therefore possible that the uranium nucleus has only small stability of form, and may, after neutron capture, divide itself into two nuclei of roughly equal size (the precise ratio of sizes depending on finer structural features and perhaps partly on chance). These two nuclei will repel each other and should gain a total kinetic energy of c. 200 Mev., as calculated from nuclear radius and charge.

So essentially, Meitner managed to explain the loss of mass observed during the nuclear reaction, as released kinetic energy. We can easily follow her reasoning by doing some basic calculations. Let's use one of the possible fission reactions:



On the left side we have the neutron that was used as a projectile to bombard the heavy element (Uranium-235). On the right side we have the resulting lighter elements that were detected, in addition to 3 neutrons.

Now, by calculating the atomic masses we have,

$$1.0086649 + 235.043928 \rightarrow 143.9229405 + 88.917630 + 3(1.0086649)$$

$$236.0525929 \rightarrow 235.8665652$$

The mass defect is then $(236.0525929 - 235.8665652) = 0.1860277\text{u}$

By introducing it (in Kg) into the Einstein's mass-energy equation, we have

$$E = (0.1860277 \times 1.66 \times 10^{-27}) \cdot c^2$$

$$E = 2.77631293 \times 10^{-11}\text{J}$$

¹ https://www.atomicarchive.com/resources/documents/beginnings/nature_meitner.html

Finally, by converting this energy to MeV we end up with approximately 173 MeV. The remaining energy, up to the “circa 200 MeV” value correctly postulated by Meitner, comes from the neutrons (~ 2 MeV per neutron), neutrinos, γ -rays and β decay products.

What was initially thought to possibly be an error in the interpretation of certain experiments, played out as one of the greatest discoveries ever made: nuclear fission.

Nuclear fission is an exothermic nuclear reaction ($Q > 0$) that releases a significant amount of energy by bombarding a fissile isotope (e.g. U-235) with a neutron. Moreover, a variable number of neutrons is also produced as a result of this nuclear reaction.

In view of that scenario, a further objective instantly emerged in the scientific community: turning this newly discovered reaction into a self-sustaining one. This could be plausibly achieved by leveraging the additional emitted neutrons to cause fission in neighboring isotopes and so on.

Obviously, in that era of global war the awareness of controlling such a powerful nuclear reaction translated into an immediate, practical use: nuclear weapons. I would like to jump over that part, in order to land on a much more useful scenario for us today: generating clean electricity.

1.2 When physics and engineering meet: Nuclear Reactors

One of the outstanding properties that emerged from the understanding of nuclear fission was that most of the released energy could be recovered through further interactions between the fission products (excluding neutrinos as it is extremely difficult for these particles to interact with others) and the matter nearby. This recovery mechanism would essentially be relying on the ability to transfer thermal energy, or heat.

So, the idea was to see if it would be possible to initiate and control a self-sustaining nuclear fission reaction inside a controlled environment, to generate thermal power in an efficient manner. If successful, this nuclear device could be used as the heat source in a thermal power plant, instead of having to burn fossil fuels.

The good news is that it was indeed possible: that device is known as a nuclear fission reactor. The bad news is that if a series of radically unexpected events and failures happen, there is a chance for this, otherwise controlled, physical process to turn into an uncontrolled scenario with the potential for serious consequences.

With that in mind, let's move to study how Pressurized Water Reactors (PWRs), the most common nuclear reactors, are designed to prevent this kind of worst-case scenario from happening.

1.3 Physics of Pressurized Water Reactors

Instead of merely presenting a description of the physical concepts behind nuclear reactors (and specifically PWRs), these will be elaborated in a step-by-step approach, following deductive reasoning when possible.

I think that this is the most fruitful way to broach such a complex topic, in order to provide the right context for those readers who do not have expertise in this field but are, in any case, familiar with certain concepts of physics and engineering.

Let's start with just three elements:

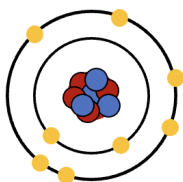
1. A hollow, finite cylindrical volume. Let's optimistically consider it a 'Reactor',
2. A neutron,
3. A fissile isotope, Uranium-235 (U-235).

As previously discussed, the main idea behind a nuclear reactor is that it should be capable of initiating and controlling a fission chain reaction. So, for now, let's place the neutron and the fissile isotope inside our 'Reactor'.

The neutron is undoubtedly one of the main characters here, so it would seem reasonable that our first step in this journey is to characterize how neutrons interact with matter.

1.3.1 - Neutron interactions with matter

The advantage of neutrons, over other subatomic particles, is that they have a neutral charge. Therefore, they are not subjected to Coulomb forces, and their interactions with electrons can be considered negligible. These properties enable neutrons to easily penetrate the atomic nucleus and interact with its nucleons, thus being a notable tool to understanding matter by analyzing the resulting interactions. On the other hand, these same properties (and others we will see shortly) work against our ability to control and 'guide' neutrons when 'controlling' matter is the objective instead.



Intuitively, we tend to visualize an atom as the image on the left, which corresponds to the Bohr model: with neutrons and protons in the nucleus and a series of electrons orbiting around them. However, when we, who are outside the world of professional physics, try to visualize the nucleus, the task is not that intuitive anymore. Analogously to Bohr's atomic model, we can find several other nuclear models (theoretical frameworks used in nuclear physics to describe the structure and behavior of the atomic nuclei).

To better understand the interactions between neutrons and matter, let's briefly discuss them.

Nuclear models can be divided into two main groups, according to the approach to study the interactions between nucleons.

1. Independent-particle models

These models focus on the interactions between individual nucleons, considering that each nucleon behaves independently, and treating the rest of the nucleons as though they were taking a passive role.

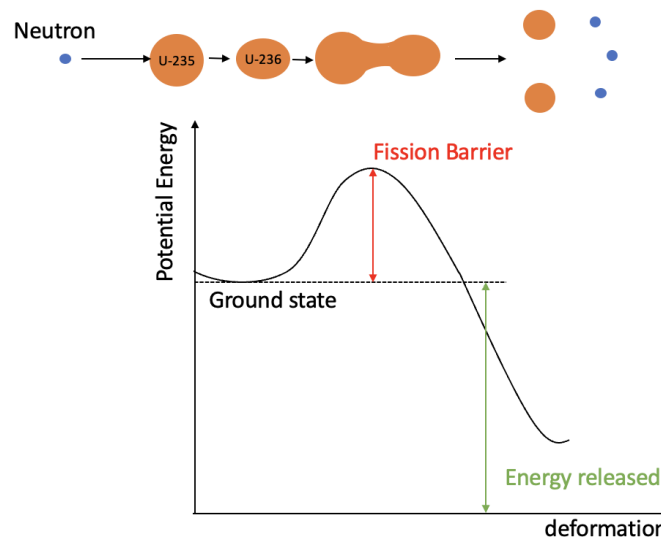
2. Strong-interaction models

For these models the main assumption is that all the nucleons behave collectively, cooperating during any potential interaction as a result of the strong nuclear force between them.

One of the interesting characteristics of the Strong-interaction models is that they describe the nucleus as a fluid. Therefore, we can think of the nucleus as a drop of nuclear liquid, similar to a water drop. Bearing this concept in mind, it will be easier to understand one of the most important interactions for a nuclear reactor: fission.

Neutron-induced Fission

In general terms, heavy isotopes can undergo spontaneous fission occasionally. However, simply waiting for that event to occur does not seem the best approach for building a nuclear reactor. Instead, the idea is to force a fission reaction by supplying the required additional energy to the target nucleus. As we have just seen, neutrons are pretty good at interacting with nuclei due to their specific characteristics, so they seem like the right candidates to supply that energy.



So, we have a neutron that will hit the U-235 nucleus, thus forming a compound nucleus (U-236), where both the kinetic and binding energy of the impinging neutron become available for the

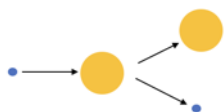
nucleons. From the perspective of the liquid-drop model, this 'traumatic' event will generate a series of elongations (imagine a water droplet 'absorbing' another water droplet), that will end up in fission when the provided energy is beyond the fission activation energy (fission barrier). After that point, the nucleus will split in two fragments, which are then accelerated by the Coulomb repulsion between them, thus gaining a significant amount of kinetic energy.

These fission products are then found in an excited state, from which they will decay as follows:

- Prompt neutrons
2 or 3 (an average number of 2.5 per fission event) prompt neutrons are emitted within 10^{-14} seconds after fission. Approximately the 99.3% of neutrons generated after fission are prompt. These neutrons will be important for sustaining the chain reaction, as we will see later.
- Gamma (γ) radiation
- Beta (β) radiation
- Delayed neutrons
After some time (from a few milliseconds to several minutes), some of the beta-decayed fission products, called precursors, will occasionally decay by emitting a delayed neutron. Although they only account for approximately 0.7% of the total neutrons generated as a result of a fission reaction, they will be crucial for the ability to control a nuclear reactor, as will be elaborated later.

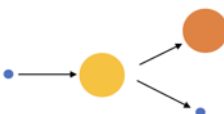
Let's now briefly describe elastic and inelastic scattering.

Elastic scattering



In this interaction, both the momentum and the kinetic energy are conserved, although there is a transfer of kinetic energy between the neutron and the target nucleus. This can be used to prevent neutrons from leaking out of the reactor, as well as to slow down them.

Inelastic scattering



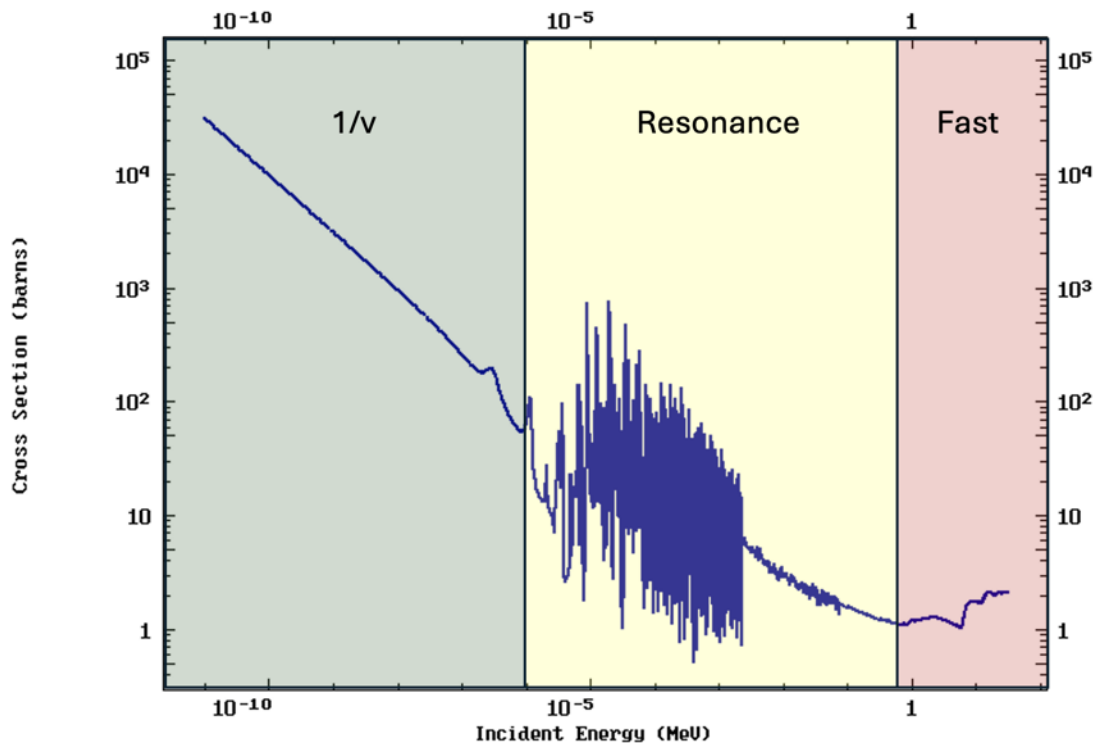
In this case, the kinetic energy of the system is not conserved because the target nucleus absorbs part of the kinetic energy of the impinging neutron to reach an excited state, from which it will later decay by emitting γ rays. Although to a lesser extent than in elastic scattering, inelastic collisions also help to slow down neutrons.

In the 'Reactivity Control' section, the reader will find a comprehensive explanation of another vital interaction for ensuring the safety and stability of nuclear reactors: resonance capture.

1.3.2 - Microscopic Cross Sections

We are assuming that the neutron and the target nucleus will interact. However, any of these interactions is inherently stochastic, so to properly characterize them we would need a measure that can represent the probability for such a potential interaction to occur. This is the idea behind microscopic cross sections.

Let's analyze the energy-dependent ($\sigma(E)$) microscopic fission cross section for U-235²

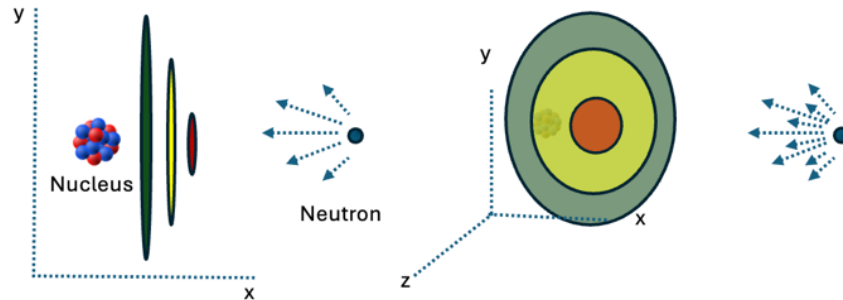


Most importantly, we can see that the microscopic cross section for a specific element (in this case U-235) has an energy-dependent behavior. There may be significant changes according to the incident energy (X axis) of the impinging neutron.

The cross section (σ) itself is measured in 'barns' (Y axis, 10^{-24}cm^2), which defines the effective area that the target nucleus presents to the neutron for a particular reaction: a large σ means more probabilities for that particular interaction to occur.

The following image depicts a simple visual representation of cross sections, just to facilitate its understanding. It should be noted that this area can be much larger (or smaller) than the geometrical cross section of the nucleus.

² <https://www-nds.iaea.org/exfor/endl.htm>



As in most heavy isotopes, the absorption (capture and fission) cross section ($\sigma_a = \sigma_c + \sigma_f$) presents three clearly defined regions:

1. Thermal

In this region, σ_f is inversely proportional to neutron velocity (which is related to its incident energy). Thus, low energy neutrons (< 1 eV) are more likely to cause fission.

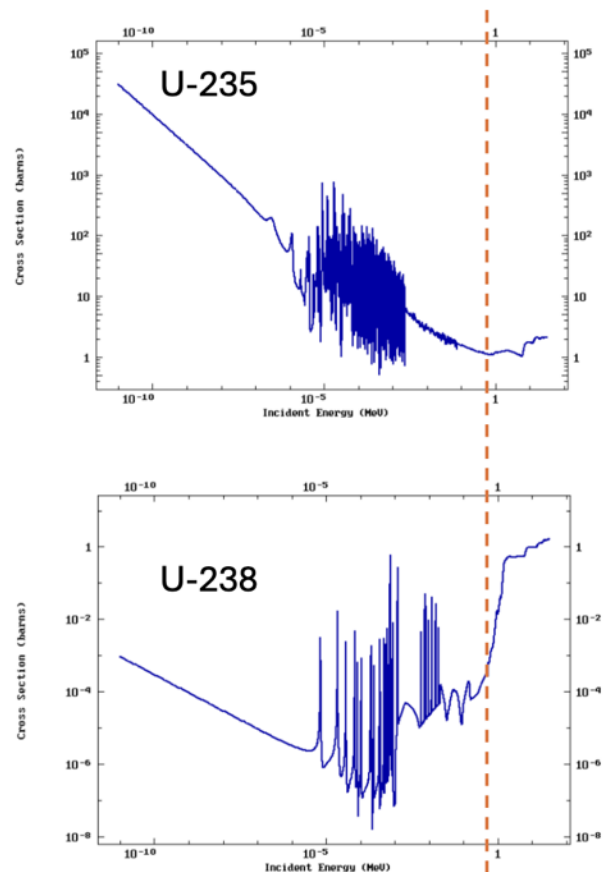
2. Resonance

In this region we find a series of resonance peaks, which are directly related to the quantum energy levels for the target compound nucleus (this is elaborated in the “Reactivity Control” section).

3. Fast

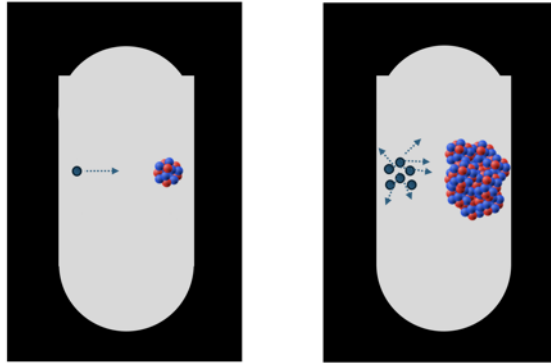
Here we need to distinguish between fissile (U-235) and fissionable (U-238) isotopes.

U-235 presents a small cross section in this area, meaning that fission-born neutrons (which are created with a kinetic energy that lies in the ‘fast’ range) don’t represent a significant contribution to cause fission. On the other hand, for U-238, it’s the opposite case, as we can see in their corresponding fission cross sections. However, U-238 has a fission cross section several orders of magnitude lower than U-235. Bear this property in mind because, as we will see later, it is crucial to the design of PWRs.



1.3.3 - From one to many

At this point, we have seen the basis of a neutron-induced fission. However, we cannot forget that the main goal in our 'Reactor' is to control and sustain a chain reaction, not just initiating it, so it is time to scale things up.



As quantities, compositions and volumes will now be important, there are some new concepts to be introduced:

1. **Instead of microscopic cross sections, now we will have macroscopic cross sections.**

The world inside our 'Reactor' will change substantially. We will find a variety of materials, composed of different nuclei in variable numbers. Thus, we cannot rely on a microscopic cross section only, instead we will have macroscopic cross sections that still represent an effective area, but this time taking into account all the nuclei present per unit volume of the target material.

Thus we have it that a macroscopic cross section of a composite material will be:

$$\Sigma = N_1 \sigma_1 + N_2 \sigma_2 + \dots N_n \sigma_n$$

Where

Σ = Macroscopic cross section (cm^{-1})

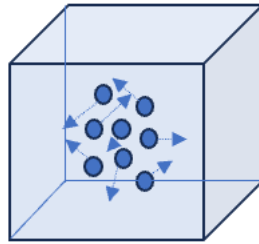
N_n = Atom density of n^{th} element (atoms/cm^3)

σ_n = Microscopic cross section (cm^2) of the n^{th} element

2. **Instead of a single neutron, we will now have a flux of neutrons.**

As we have seen, the macroscopic cross sections will give us a measurement of the probability of a neutron undergoing a specific reaction per centimeter of travel through the material. However, we're leaving behind the single-neutron (or uniform neutron beams) 'world', so we will need to adapt our model to this change as well.

To approach this task, we will take into account the density of neutrons per cm^3 and their velocities. As previously mentioned, neutrons are difficult to guide so we have to assume that the neutrons inside our 'Reactor' will be moving in different directions. As these directions do not significantly impact the potential interaction between the impinging neutron and the nucleus, the total flux will be the scalar sum of the intensity (neutron density times average velocity) of these different neutron beams in unit volume per unit time.



Therefore, we can think of this flux as the total path length covered by all neutrons in one cm^3 during one second. So, we have $\Phi = n v$

Where

Φ = neutron flux in neutrons per cm^2 in a second ($\text{n.cm}^{-2}.\text{s}^{-1}$)

n = neutron density (neutrons/ cm^3)

v = neutron velocity (cm.s^{-1})

3. Instead of a single isotope of U-235 we will now introduce natural uranium.

Nuclear fuel is obviously a fundamental part of the reactor. For now, let's place some natural uranium (0.711% of U-235 and 99.284% of U-238). Spoiler: this composition will not last long.

To sum up the situation in our 'Reactor' after these changes: we have a neutron flux ready to hit natural uranium. To make things even more ideal in our custom 'Reactor', we can control time so let's assume everything is on 'pause' now. What would happen if we pressed 'play'?

According to what we have seen, the neutron flux will most probably be able to trigger a fission reaction in the fissile part of the nuclear fuel (U-235), which in turn will generate 3 (rounding it to make things easier) prompt neutrons, and occasionally a delayed one.

Solely based on the previous assumption, we can now consider our 'Reactor' as a neutron multiplication device. As we can press 'pause' and 'play' at will, we can start taking measurements between fission events, thus formulating a multiplication factor (k), which would define the ratio at which the neutron population changes.

$$k = \frac{\text{Number of neutrons in one generation}}{\text{Number of neutrons in the previous generation}}$$

Based on this value we will initially define three states for our 'Reactor':

1. Supercritical ($k > 1$)

As stated in the definition of neutron flux, if the number of neutrons has increased, then neutron flux is also increasing. This yields more neutrons than those required for a self-sustaining chain reaction. Therefore, during this state large amounts of energy will be generated.

2. Critical ($k = 1$)

The neutron population remains the same over different generations, thus we shouldn't expect significant changes in the neutron flux nor in the energy generated. The self-sustaining chain reaction is maintained.

3. Subcritical ($k < 1$)

As the number of neutrons available for producing fission is decreasing, the self-sustaining chain reaction cannot be maintained, thus limiting the production of energy.

Based on the 'Criticality' of the reactor we will also define 'Reactivity' (ρ) as a parameter that denotes a deviation from the 'Critical' state ($k = 1$, $\rho = 0$). As a result, a positive reactivity means that the reactor is increasing the neutron population. On the other hand, a negative reactivity provides a decreasing neutron population.

$$\rho = (k - 1)/k$$

Unfortunately, in the real world we cannot 'pause' and 'play' the reactor so easily, so if we are planning to control a fission chain reaction, there is a fundamental parameter that comes into play: time.

1.3.4 - Time matters

Fission is a fast phenomenon: prompt neutrons are generated within 10^{-14} s after fission. Assuming each of these prompt neutrons in turn generates further fissions in other nuclei, we will have a brutal exponential growth in a fraction of a second. We can see an example of this by removing from the following point kinetics differential equation the contribution (red-highlighted terms) of the delayed neutrons.

$$\frac{dn}{dt} = \frac{k(1 - \beta) - 1}{\ell} + \sum_i \lambda_i C_i$$

and resolving

$$n(t) = n_0 e^{\frac{\Delta k_{eff}}{\ell} t}$$

Where

- Δk_{eff} is the increment in the multiplication factor (constant)
- n_0 is the initial density of neutrons (1)
- ℓ is the prompt neutron lifetime ($\sim 10^{-4}$ s in a regular thermal reactor)
- t is the time (1s)

By relying exclusively on the prompt neutrons, for a $\Delta k_{eff} = 0.001$ in just one second, we would be dealing with a neutron population of $\sim 22k$, generated over 10 reactor periods (e^{10}).

Considering that the power level of a reactor is proportional to its neutron population (or neutron flux), it seems clear that unless you want to build a nuclear weapon, or replicate Chernobyl, this scenario poses a serious problem: the reactor could become “prompt critical”³, making it impossible to design a control system intended to deal with such conditions.

Fortunately, delayed neutrons can be used to overcome this issue, The reactor is initially allowed to become sub-critical on prompt neutrons and then delayed neutrons are utilized to reach criticality. By using this approach, we guarantee that the multiplication factor will depend on the generation time of the delayed neutrons, thus increasing the expected available time for the industrial control system to respond.

³ https://en.wikipedia.org/wiki/Prompt_criticality

1.3.5 - Slow down, please.

Let's recap the constraints that apply to our 'Reactor' so far, in known terms:

- We are currently using natural Uranium as nuclear fuel, whose composition is ~99.3% U-238 and ~0.7% U-235.
 - U-238 cannot sustain a chain reaction.
 - It is fissionable.
 - Neutrons generated during a U-238 fission reaction have less energy than required to generate further fissions in the U-238.
 - On average, it brings a smaller number of neutrons generated per fission compared with U-235.
 - U-238 fission cross section for fast neutrons is several orders of magnitude smaller than the fission cross section of U-235 for thermal neutrons.
 - U-238 absorption (resonance capture) cross section is large.
 - U-235 can sustain a chain reaction.
 - U-235 fission cross section is large for thermal neutrons.
 - U-235 fission cross section is not significant for fast neutrons.
 - It is fissile.
- We need to design a control system for the reactor, with a reasonable response time.
 - We cannot rely on prompt neutrons only to sustain a chain reaction (too fast to handle).
 - We need to leverage delayed neutrons.
 - Delayed neutrons have lower energies than prompt neutrons.

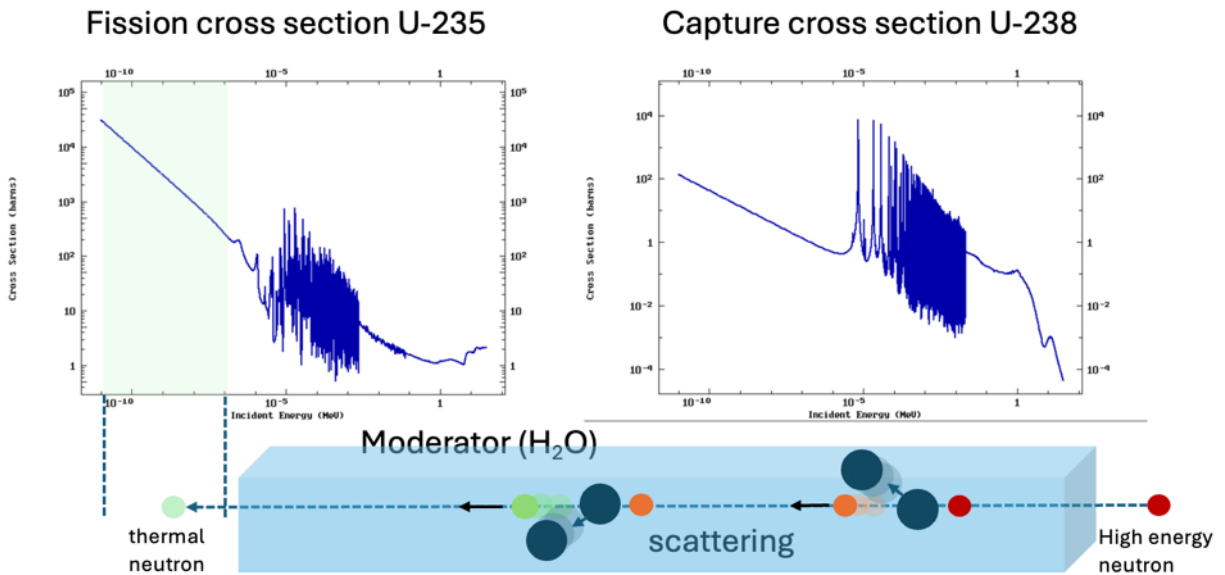
Based on these premises, if we want to move towards a self-sustaining chain reaction we will need to perform two actions:

1. Slowing down fast neutrons to be able to sustain a chain reaction

It's time to introduce another element into our 'Reactor': the moderator.

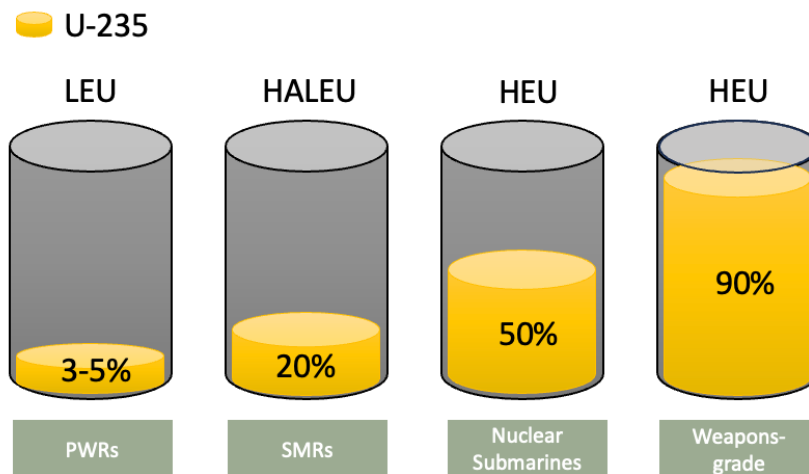
As we elaborated in the section 'Neutron interactions with matter', during elastic (and occasionally inelastic) scattering, the neutrons can lose part of their energy. The moderator material is basically in charge of optimizing this interaction by slowing down neutrons until they reach their thermal range (0.01 to 0.1 eV) in the most efficient way. Additionally, as the neutrons are interacting with the moderator, they are kept away from the U-238 present in the fuel (in a heterogeneous reactor, where fuel and moderator are separated) thus avoiding its large resonance capture area. When these neutrons diffuse back to the fuel, already thermalized, they will be able to cause further fissions.

It turns out that, due to its low mass, large scattering cross section and low cost (compared with heavy water), a common element we all know is the perfect moderator: water.



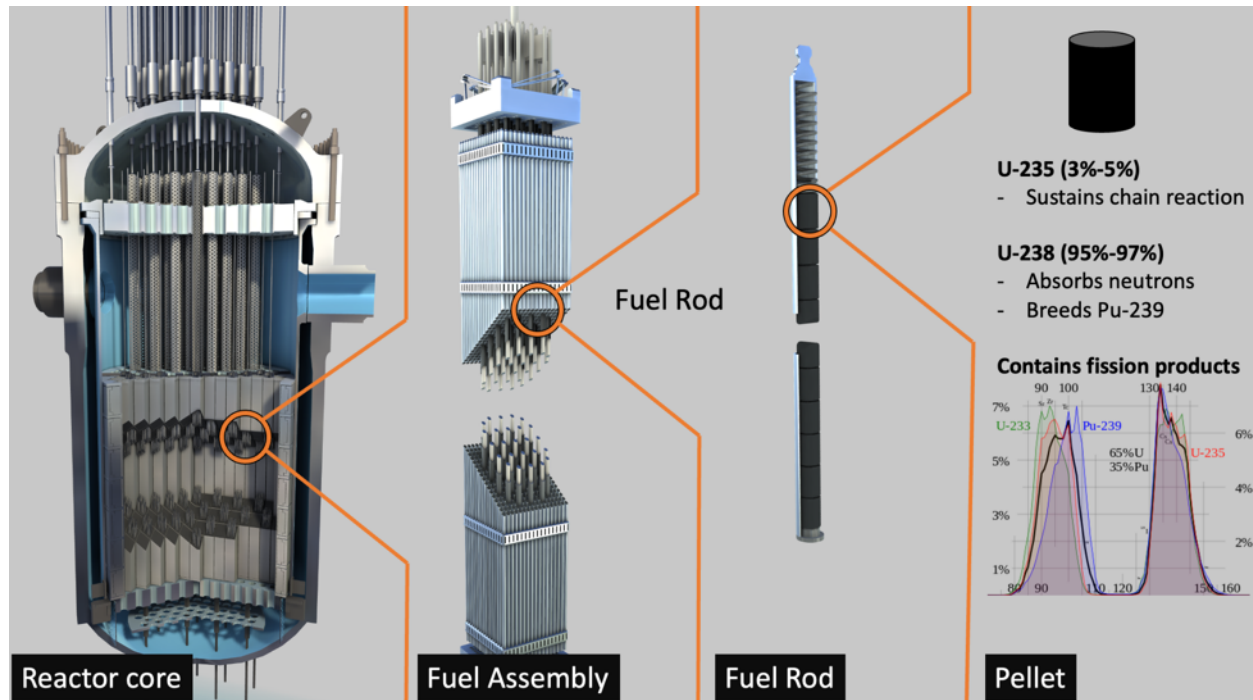
2. Increasing the amount of U-235 in the nuclear fuel

Uranium enrichment is surely a familiar term for the reader (e.g. Stuxnet). When natural uranium is enriched, it means that its concentration of U-235, the fissile isotope, is increased to ensure the ability to achieve a self-sustaining chain reaction under different conditions. Thus, we have highly enriched uranium (HEU > 20%) that can be used either for nuclear weapons (weapons-grade HEU is approximately 90% enriched to make up for the use of fast neutrons) or low-enriched uranium (LEU, between 3% and 5%, increasing from just 0.7% present in natural uranium) and High-Assay Low-Enriched Uranium (HALEU) for nuclear fuel.



1.3.6 - Nuclear fuel

Having two new elements, the moderator and the new composition (low-enriched uranium) of the nuclear fuel, we should now see how they fit together.



4

PWR are heterogeneous reactors, where the moderator is separated from the nuclear fuel, which is lumped in the form of pellets. These pellets are mainly composed of UO_2 , (other compositions are possible, e.g. MOX fuel⁵), but may contain other elements such as gadolinium (a neutron poison we'll talk about later).

A regular nuclear fuel rod consists of several pellets arranged inside zircaloy tubes (cladding).

Inside a fuel rod, the thin gap between pellets and cladding is filled with helium gas to enhance heat transfer to the coolant. This gap also accommodates the expected buildup of fission products as fuel is burnt up.

These fuel rods are then arranged in fuel assemblies, with their number and geometry varying according to different reactor designs.

⁴ Certain elements of this image have been obtained from <https://www.energyencyclopedia.com/en/free-downloads/images>

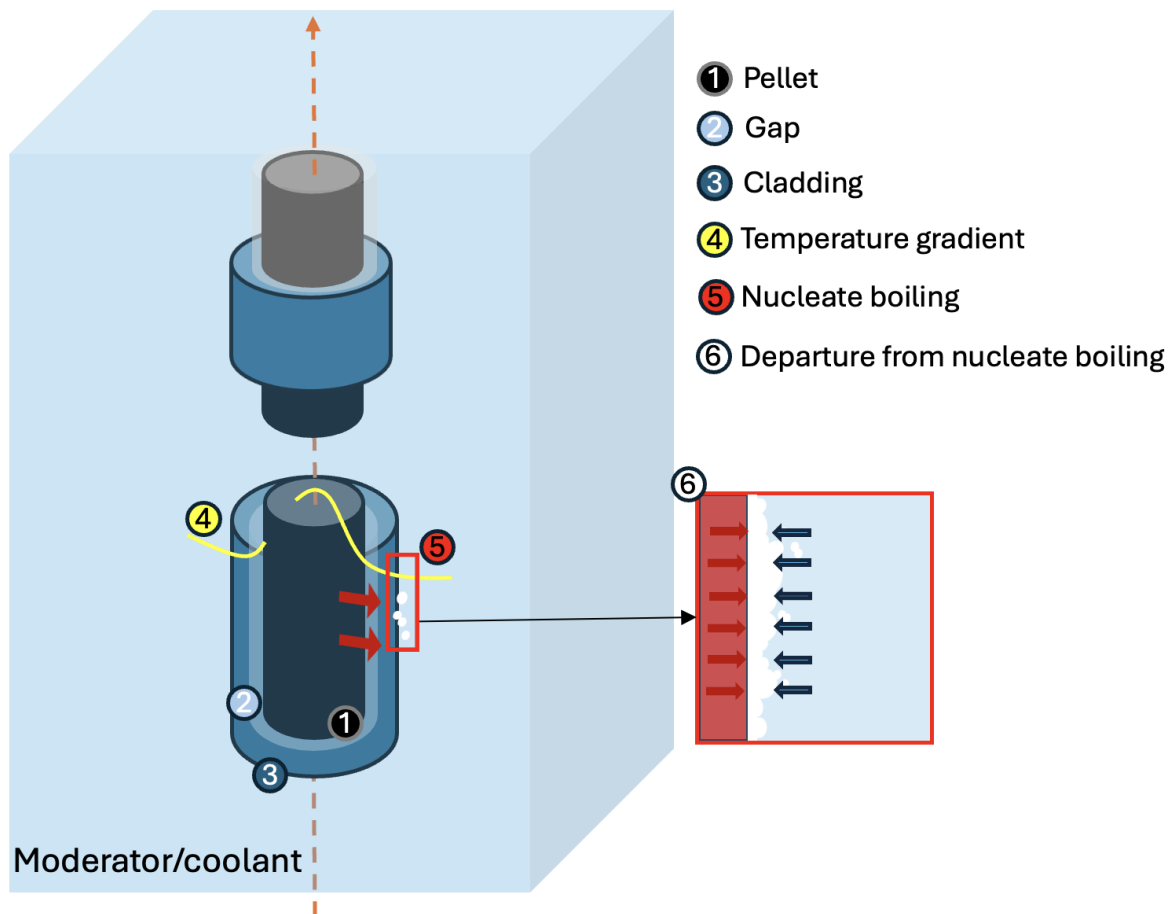
⁵ https://en.wikipedia.org/wiki/MOX_fuel

Basically, fission is happening in the pellets, thus generating a large amount of thermal energy (heat). As UO_2 performs extremely poorly at conducting heat, a temperature gradient between the fuel centerline and the moderator/coolant allows this thermal energy to be transferred to the latter via conduction. As in a PWR the moderator is also the coolant (water), this is the essential mechanism used to remove, and recover, from the core the heat generated by nuclear fission.

The water is heated as a result of this interaction and used to boil additional water in the steam generators. The high-quality steam generated is used to drive a turbine that will generate the electricity that is injected into the grid. This is how a nuclear power plant works in a simplistic manner, although we'll elaborate this mechanism later.

In the interest of introducing the cyber-physical component of the research, I want to acquaint you with a singularly important phenomenon which significantly affects the stability of the reactor: Departure from Nucleate boiling.

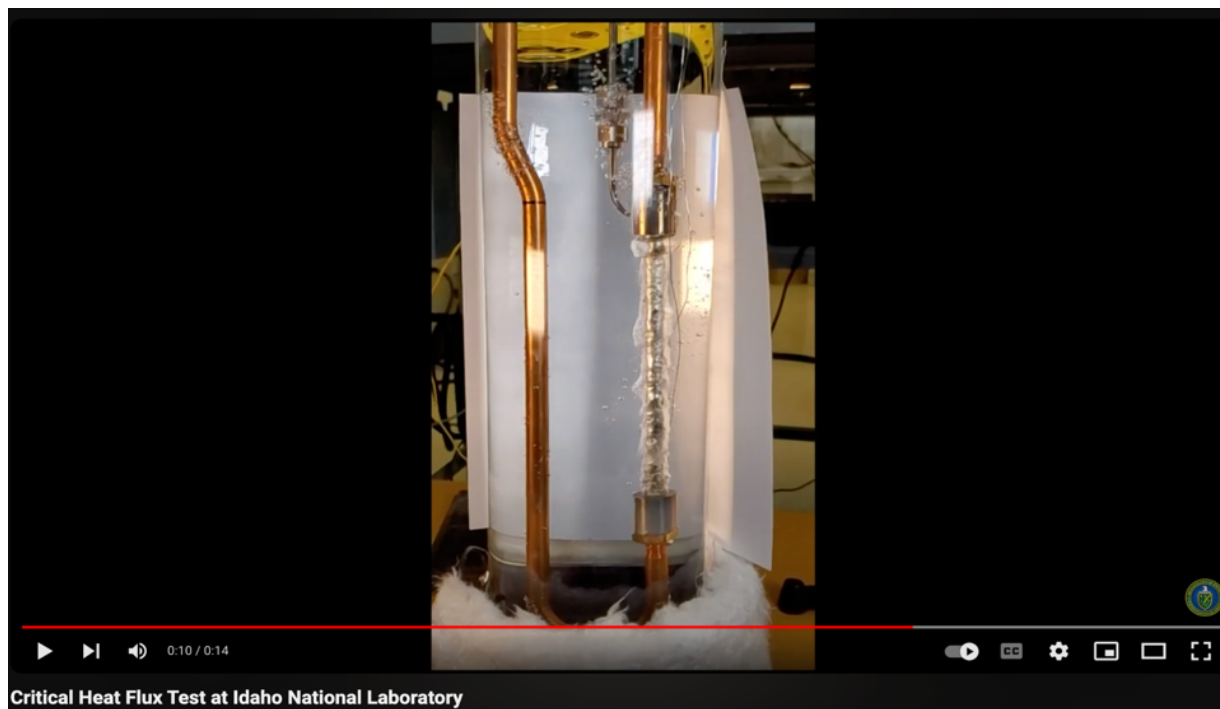
Departure from Nucleate Boiling.



The thermodynamic interactions on the surface between the coolant and the cladding are subjected to different regimes (boiling curve⁶), depending on the amount of heat that is being added to this system.

Below the Critical Heat Flux (CHF), the gas and liquid phases are at 'equilibrium', in what is called 'nucleate boiling'. Basically, a certain amount of steam, in the form of bubbles, develops due to irregularities in the microgeometry of the cladding surface. These hot bubbles will eventually collapse in the coolant, thus increasing the heat transfer from the cladding surface to the coolant. You can look inside a kettle, and you'll see something similar.

However, above the CHF, things can go wrong fast. The departure from nucleate boiling occurs when enough heat is added to the system, thus leading to the formation of a film of steam around the surface, as we can see in the diagram above. As opposed to the bubbles, this newly formed gaseous layer doesn't collapse because more heat is being generated than that which can be removed by the coolant from the cladding surface. This effectively degrades the heat transfer coefficient between the nuclear fuel and the coolant, increasing the fuel temperature.



7

As the heat cannot be removed, the excessive fuel temperature may result in the loss of structural integrity of the pellets, resulting in the release of fission products and radioactivity. If this situation is not addressed, the temperature in the fuel will keep increasing, eventually causing a core melt.

⁶ https://en.wikipedia.org/wiki/Nucleate_boiling

⁷ https://www.youtube.com/watch?v=5xB_gLfdJM

1.3.7 - Reactivity Control

For a reactor to be successful it's not enough to trigger a chain reaction, we need to control it as well. Therefore, to keep our reactor stable, we need ways to predict its criticality, in order to avoid uncontrolled power excursions that may ultimately damage the nuclear fuel. We previously established reactivity as a deviation from a critical reactor, so it would make sense to assume that reactivity control is the key to designing a stable reactor.

A nuclear reactor is probably the worst place to leave something to chance. The best way to ensure safety and stability is to perform a detailed analysis of everything that may realistically happen, either under normal conditions or when something goes wrong. Among all the different analyses required to accomplish this task, one of the most important aspects of designing a reactor core is characterizing reactivity changes. This means that we are now entering into the land of Reactor Dynamics, a wide and complex area. Let's focus on the most important concepts required to comprehend the subsequent 'cyber-physical attacks' section.

The neutrons generated inside the reactor will have different 'fates', so the six-factor formula is a mathematical way to quantify them. Essentially this formula covers the different events (physical processes that a fission-induced neutron undergoes during a generation) in a neutron's life cycle to determine the resulting effective multiplication factor (k_{eff}), a vital property to study reactor control.

$$K_{eff} = \varepsilon \cdot P_f \cdot P_{th} \cdot \rho \cdot f \cdot \eta$$

Each of these factors can be used to assess how reactivity can be controlled by examining the different variables that influence them.

Where:

1. Fast Fission (ε)

This is the ratio of the net number of fast neutrons produced, regardless of their energy, to the fast neutrons produced due to thermal fissions.

$$\varepsilon = \frac{\text{number of fast neutrons produced by fissions at all energies}}{\text{number of fast neutrons produced in thermal fission}}$$

The most significant variables that impact this factor would be:

- a) The disposition of the core elements (e.g. fuel rods)

In a heterogeneous thermal reactor, the generated fast neutrons will interact with the moderator that surrounds the fuel, however the closer the fuel rods are placed,

the fewer opportunities for the fast neutrons to be thermalized before diffusing back to the fuel.

For instance, nuclear bombs fundamentally depend on fast fission. So, to compensate for the low probability of fast neutrons causing fission in U-235, HEU (~90%) is used and placed in a very compact volume.

b) The concentration of U-238

In the nuclear fuel there is a significant concentration of U-238 (around 95%-97.5%). So, although low, the possibility of fast neutrons causing fission in U-238 still needs to be considered.

It is worth mentioning that core temperature may slightly increase ϵ . When the temperature of the core increases, the density of the water (moderator) decreases, thus losing part of its moderating ability. This will increase the number of fast fission neutrons able to diffuse back to the fuel while keeping high energies, due to the reduced moderator-to-fuel ratio.

2. Fast non-leakage (P_f)

A nuclear reactor is a finite volume, so neutrons can leak out. The non-leakage event is divided in two terms, P_{th} (thermal) and P_f (although they could be combined), according to the energy of the neutron. P_f (fast) quantifies the probability that a neutron does not leak out while it is still slowing down from high energies. Remember that fission-induced neutrons are generated at high energies, which means its mean free path is significant.

$$P_f = \frac{\text{number of fast neutrons that do not leak from the reactor core during the slowing down process}}{\text{number of fast neutrons produced by fissions at all energies}}$$

Core temperature is a fundamental variable here, if it increases, the moderator density will decrease. This means that neutrons will have fewer nuclei to scatter with, thus finding a small number of 'obstacles' on their way out of the reactor. As a result, the higher the temperature, the lower the P_f .

3. Thermal non-leakage (P_{th})

$$P_{th} = \frac{\text{number of thermal neutrons that do not leak from the reactor core during the neutron diffusion process}}{\text{number of neutrons that reach thermal energies}}$$

Instead of taking account of fast neutrons as in the previous factor, P_{th} quantifies those neutrons already thermalized. However, the effect that the temperature of the core has in this factor is the same as in P_f .

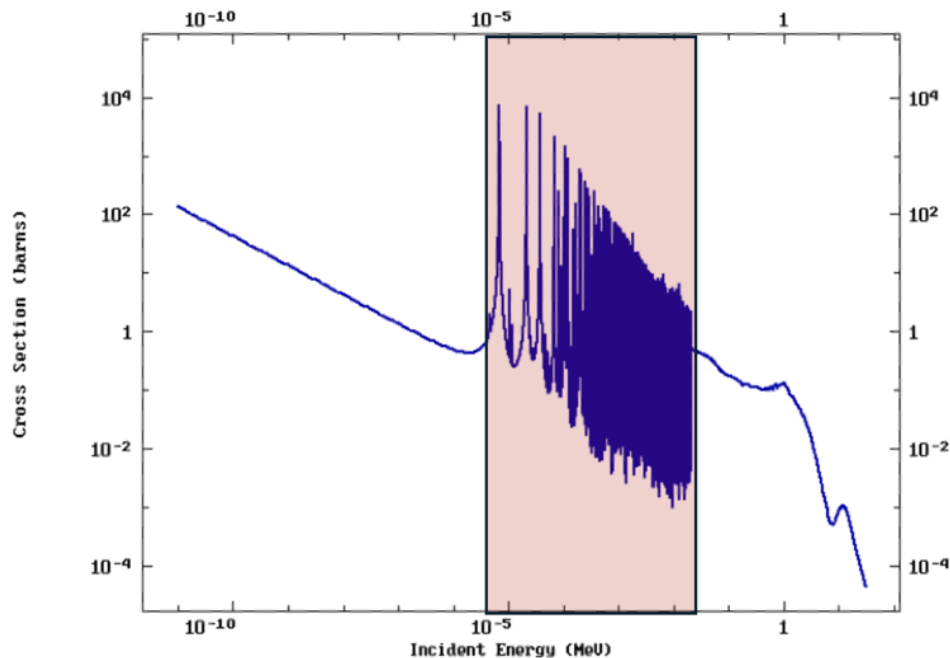
Another contributing factor in this case is the fuel burnup, as thermalized neutrons will have more difficulty causing further fissions due to the decreasing amount of fissile material. As a result, P_{th} will also decrease during the fuel cycle.

4. Resonance escape probability (ρ)

All factors are important, but this one is especially interesting and key for the safety of modern reactors. The interpretation is straightforward: the probability that a fast neutron reaches thermal energies without being captured (absorbed without inducing fission). However, its analysis is quite complex, with multiple ramifications.

$$\rho = \frac{\text{number of neutrons that reach thermal energies}}{\text{number of fast neutrons that start to slow down}}$$

Fast neutrons don't have a quiet path towards reaching thermal energies. Among the challenges they face, some of which have been outlined in the previous factors, there is a truly complicated one: the resonance region in the (non-fission) capture cross-section of U-238.



For fast neutrons to have a chance of being thermalized, thus causing further fissions, they first need to escape from this large region. And it's not easy.

There are several phenomena and characteristics that influence ρ :

a) Core geometry

As in the Fast-fission factor (ϵ), the arrangement of the fuel elements and the moderator greatly impacts ρ . Let's remember that one of the functions of the moderator in a heterogeneous thermal reactor is to keep fast neutrons 'away' from U-238 while they are being thermalized, precisely to avoid being captured.

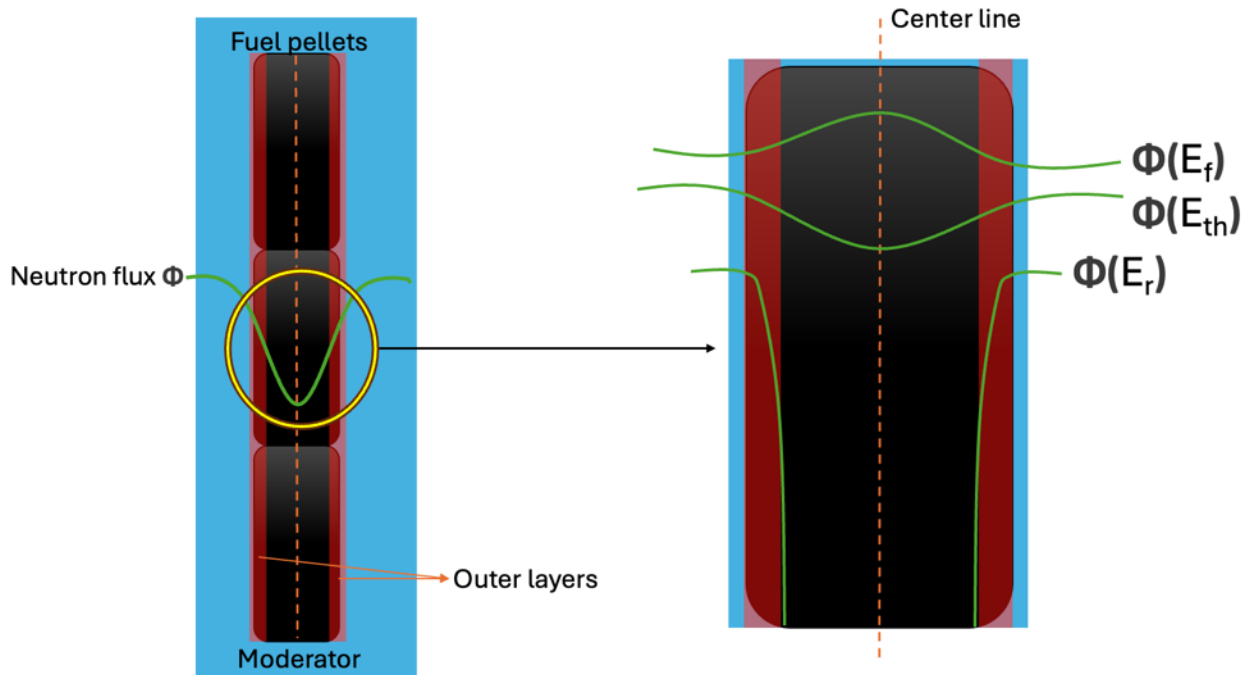
b) Spatial and energy self-shielding

Without entering into the study of neutron diffusion equations, which would require writing an entire separate book, there are some important concepts that we need to bear in mind to properly understand both energy and spatial self-shielding.

- The neutron flux in a nuclear thermal reactor is not monoenergetic. Instead, neutrons have a wide energy spectrum.
- As we have seen, cross-sections are not constant, but energy dependent.
- Although with a marginal impact compared to the actual moderator (H_2O), the different nuclei that are found inside the fuel pellets can also act as a 'moderator' as neutrons may scatter against them. As a result, the volume of the pellets is important.

Now, let's visualize these concepts. On the left side of the diagram below we have the following items:

- Fuel pellets
Nuclear fuel pellets that are placed inside the fuel rods (cladding is obviated in the diagram).
- Moderator/Coolant
The moderator is separated from the fuel, surrounding it.
- Outer layers
These represent the external surface of the pellet.
- Neutron flux Φ
We can observe a profound depression of the flux as it passes through the pellets, starting in the vicinity of the moderator-fuel area of contact.



On the right-hand side, a more detailed view of the interaction between the flux Φ and the pellet is depicted.

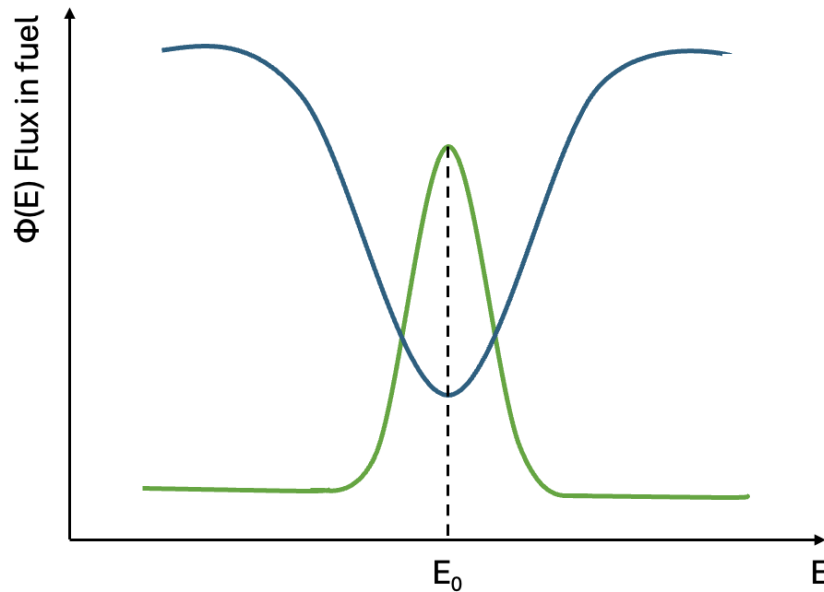
We have a decomposition of the flux into different groups, according to some of the energy ranges we can identify in the spectrum (just a sample of them for ease of explanation)

- **$\Phi(E_f)$ - Fission**
The flux of neutrons with fission energies presents a peak around the center line. The reason is that fission-induced neutrons are created with fission energies inside the fuel, then 'escape' to the moderator.
- **$\Phi(E_{th})$ - Thermal**
For thermal neutrons, we observe a slightly depressed flux, as a certain number will induce fissions in the outer layers.
- **$\Phi(E_r)$ - Resonance**
Those neutrons with resonance energies diffusing back into the fuel from the moderator will be mainly captured (due to the large U-238 resonance area) in the outer layers.

The profound depression of $\Phi(E_r)$ means that the inside part of the fuel pellet will see a significant decrease in the number of resonance neutrons. As a result, fewer neutrons will be absorbed in the resonance area of the U-238, thus increasing the

resonance escape probability (ρ). This phenomenon is known as 'spatial self-shielding', and basically it allows a self-sustaining chain reaction by using lumped fuel.

Analogously, we have the energy self-shielding phenomenon. In the simple diagram below we have the following items:



- The green curve represents a resonance region of a nucleus from the fuel, where E_0 represents the resonance peak.
- The blue curve represents the neutron flux passing through the fuel.

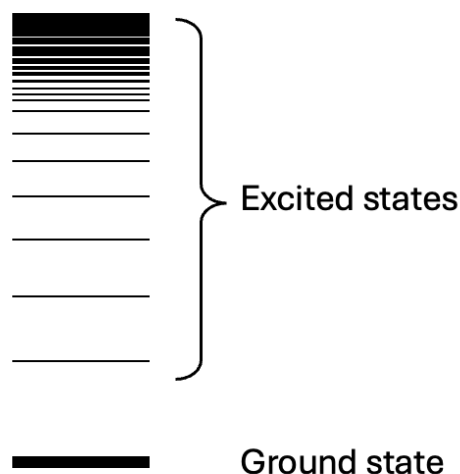
We have a depression of the flux in the vicinity of the resonance peak, similar to what we observed in the vicinity of the nuclear pellet. The result is also similar: neutrons impinging the nuclei with energies around E_0 will most probably be absorbed, thus shielding the absorber nuclei. That's why we have the flux depression.

Energy self-shielding is a fundamental piece to understanding one of the most important, and beautiful, phenomena that acts as a pillar for the safety of nuclear reactors: Doppler-broadening of resonances.

c) Doppler-broadening of resonances

We have been talking about resonances and neutrons being captured, but to better understand these concepts, let's dig a bit further into how things work at the nucleus level, according to quantum mechanics.

All nuclei have a series of different discrete energy levels: the ground state, where the nucleus is stable, and then a variable number of excited states which are defined by the nucleus' internal structure.



When a neutron comes close enough to the U-238 present in the fuel, the strong nuclear force causes them to combine, resulting in a compound nucleus reaction where the neutron's energy is rearranged among the existing nucleons. Essentially, the binding and kinetic energy of the neutron become available for the compound nucleus ($\text{U-238} + n \rightarrow \text{U-239}$).

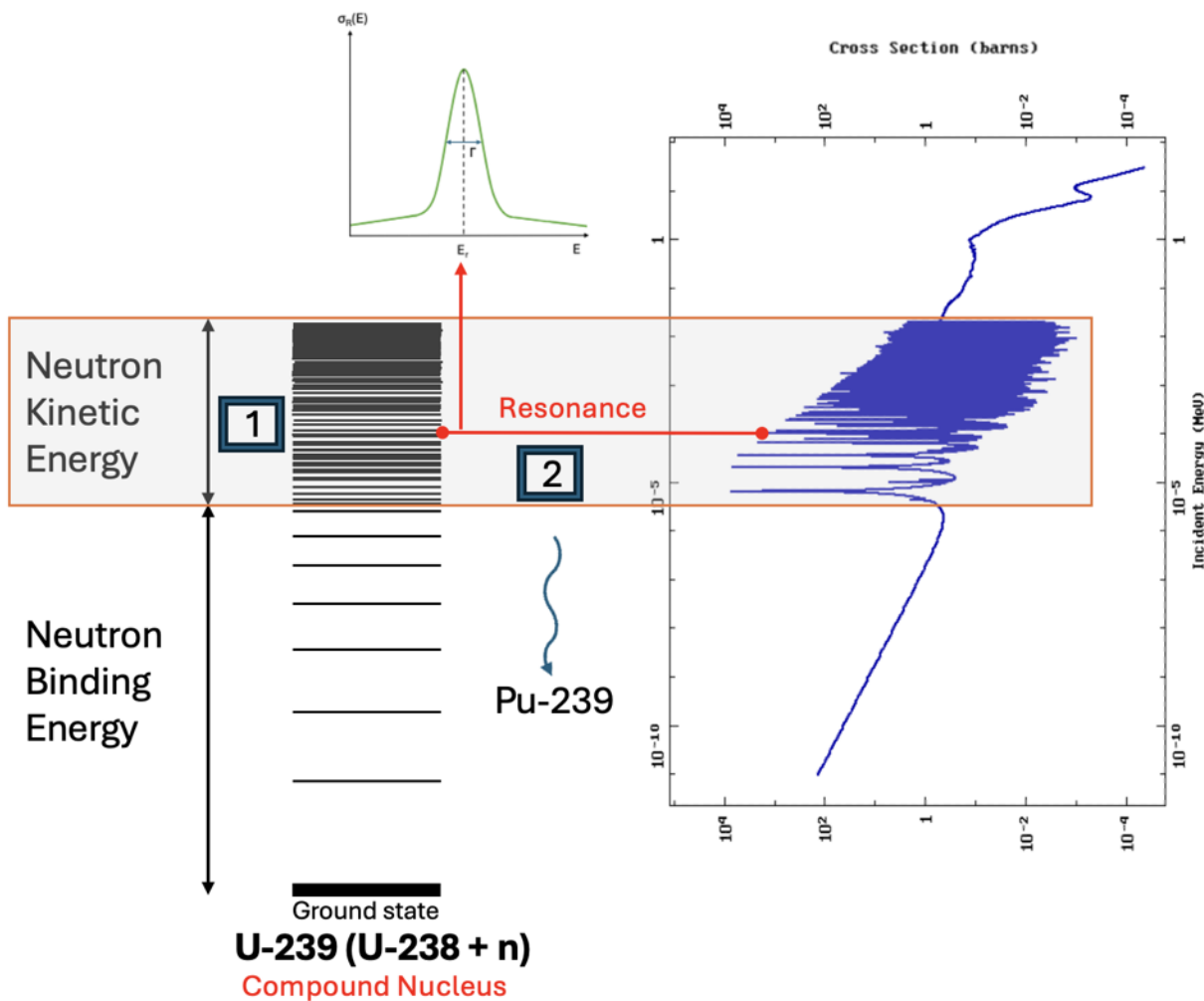
As we can see in the diagram⁸ below:

1. In the compound nucleus the density of energy levels (excited states) above the neutron's binding energy is significant. We can relate these energy levels, the neutron's energy and the 'famous' large resonance region in the capture cross-section of the U-238 previously laid out.
2. When the kinetic energy of the impinging neutron matches any of those excited states (resonances) in that region, the resulting compound nucleus (U-239) will end up beta-decaying into a new fissile isotope: Plutonium-239.

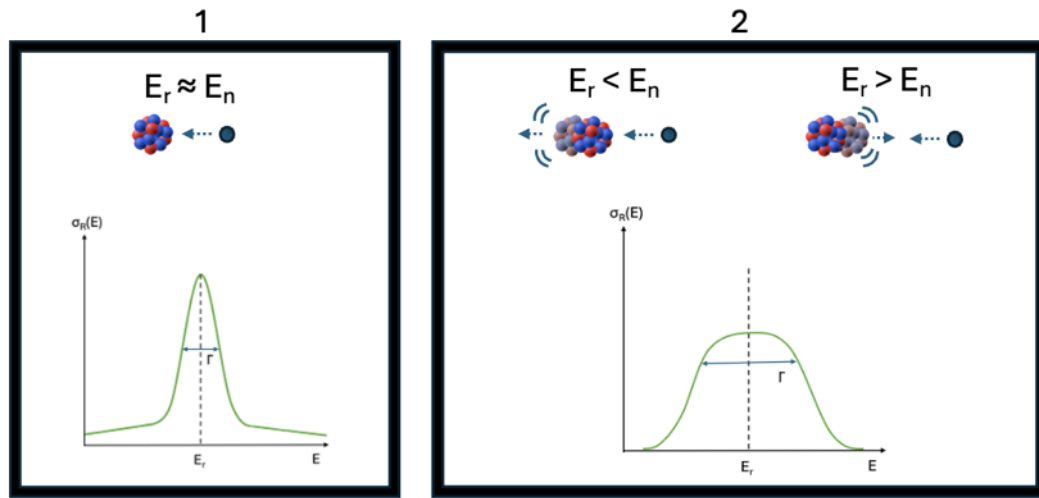
⁸ It doesn't depict an accurate representation of the different energy levels but an approximation to clarify the explanation.

This is also the reason why nuclear reactors can be used to breed weapons-grade fissile materials (Plutonium). However, these highly irradiated compounds are not as stable as HEU, which is still the preferred material to build nuclear bombs.

Essentially, this is what happens behind the scenes when a neutron cannot escape the U-238 resonance and is subsequently captured.



We have seen that the kinetic energy of the neutron is a key factor. However, it is crucial to clarify that the speed of the impinging neutron (and therefore its kinetic energy) is relative to the U-238 nucleus. As a result, when the temperature of the fuel increases, the U-238 nuclei will increase their vibration, due to thermal motion, so we have two different scenarios, as depicted in the following diagram:



1. 'Regular' resonances

This case represents the neutron-nucleus interaction assuming the U-238 nucleus is 'at rest' (not very common). Thus, to be absorbed, its kinetic energy should exactly match one of the excited states in the compound nucleus.

2. Doppler-broadened resonances

The U-238 nucleus is vibrating, thus broadening the range of the neutron energies that will make it suitable to be absorbed. If the neutron is moving in the same direction as the U-238 nucleus ($E_r < E_n$), the kinetic energy of the neutron may be higher than the one required to match the resonance (E_r), but the vibration of the U-238 nucleus makes up for that difference. The same logic applies for the opposite case ($E_r > E_n$), when the neutron and the U-238 nucleus are moving in opposite directions, and E_n is lower than the energy required to hit the resonance.

Ultimately, this means that as the fuel temperature increases, more neutrons will be absorbed. We will see later, when studying reactivity feedback mechanisms, how this effect plays a pivotal role in the safety of nuclear reactors.

5. Thermal utilization (f)

Assuming that a certain number of neutrons managed to escape from the resonance region previously shown, this factor represents the ratio of the probability of such a neutron being thermalized eventually and absorbed in the fuel to the probability that it is absorbed elsewhere in the reactor core.

$$f = \frac{\text{number of thermal neutrons absorbed in the nuclear fuel}}{\text{number of neutrons absorbed in all the material that makes up the core}}$$

In this case, this factor can be greatly influenced directly by the operator, by introducing elements in the reactor core that can be used to absorb neutrons, such as control rods, burnable poisons or soluble ones such as boric acid.

6. Reproduction (η)

This is the average number of fission neutrons produced per absorption of a thermal neutron in the fuel. Only fissile elements are considered, U-235 and Pu-239, so η is essentially constant during the fuel cycle as the amount of Pu-239 increases due to fuel burnup and U-235 decreases for the same reason.

$$\eta = \frac{\text{number of fast neutrons produced by thermal fission}}{\text{number of thermal neutrons absorbed in the nuclear fuel}}$$

η can also be influenced by the 'uncontrollable' poisons, when there are significant variations in the neutron flux, mainly Xenon-135 and to a lesser extent Samarium-139.

By studying how these factors can be used to adjust k_{eff} it is possible to determine two main schemas to control reactivity.

1. Externally

It should be possible for the nuclear operators to control reactivity directly, in order to properly handle the different conditions under which the reactor may need to operate. There is a fundamental reactor characteristic that has the potential to allow this behavior: composition

By inserting specific materials into the reactor core, it is possible to change its composition. The main objective of these materials is to control the neutron population via different mechanisms, thus ultimately controlling the reactivity.

In general terms, these materials (also called 'poisons') will be characterized by a large absorption cross-section, which allows them to neutralize reactivity excursions by absorbing neutrons.

Depending on the requirements of these materials we can determine three main purposes:

a) Shutdown control

A reactor trip, or scram, is a safety action intended to rapidly add enough negative reactivity to be able to stop the self-sustaining chain reaction.

This is an immediate operation that can be triggered either manual or automatically, requiring a high reliability within a tight timeframe.

b) Shim control

When a nuclear reactor is loaded with fresh fuel, there is an excess reactivity intended to compensate for the long-term fuel depletion due to burnup, as well as the fission products buildup.

This excess reactivity needs to be controlled, especially during the beginning of the fuel cycle. We can consider this as a medium to long-term operation.

c) Power regulation

Operators may need to adjust the reactor power level according to electric grid demand (e.g. NPP in load-following mode), or due to other operational requirements of the plant (e.g. coastdown, stretchout). This kind of reactivity control usually needs to address small reactivity transients, and anticipated situations, in the short-term.

Additionally, there are three main approaches to implement the previous reactivity control operations:

1. Movable

Movable control rods are a cluster of devices, containing neutron absorber elements, intended to be inserted (\downarrow reactivity) into, or withdrawn (\uparrow reactivity) from, the fuel assembly; either manually by the operator or automatically by the reactor protection system.

These control rods can be used in the three scenarios previously described.

2. Soluble

Operators can decide to add into the moderator a certain amount of neutron-absorber elements, such as boric acid, through the Chemical and Volume Control system. This is usually implemented in the 'Shim Control' scenario.

3. Burnable

This mechanism may overlap with the 'Movable' one (you can use burnable poisons in the control rods), but it is worth clarifying that it may or may not be under the dynamic

control of the operators: for instance, certain pellets inside the fuel assemblies may be covered, or synthesized, with a specific amount of burnable neutron poisons (e.g. gadolinium).

These burnable poisons are used for 'Shim Control' to help extend the life of a fuel cycle as well as to control excessive reactivity, especially when fresh fuel is loaded. As the fuel depletes, these poisons are also burnt up, so they progressively lose their ability to absorb neutrons, thus helping to balance the required reactivity.

There are also certain burnable poisons that are inherently generated by fission products, such as Samarium or Xenon. The cycle of these poisons needs to be considered when operating the reactor to ensure they are properly neutralized. One of the circumstances which contributed to the Chernobyl disaster was a Xenon poisoning.

2. Inherently

Reactivity ultimately depends on the macroscopic cross-sections, which in turn depend on the atomic number densities (N) of the different elements that compose the material (remember the definition of a macroscopic cross-sections $\Sigma = N_1 \sigma_1 + N_2 \sigma_2 + \dots N_n \sigma_n$). During the analysis of the six-factor formula we have seen how the density of the different core materials (fuel, moderator, etc.) is closely related to the reactor power level through a main variable: temperature.

A change in the temperature will affect the density of the macroscopic cross-sections, thus shaping the neutron flux, which will ultimately control the reactor power level. As a result, nuclear reactors use this variation of reactivity according to the temperature as a feedback mechanism to inherently sustain the stability of the reactor.

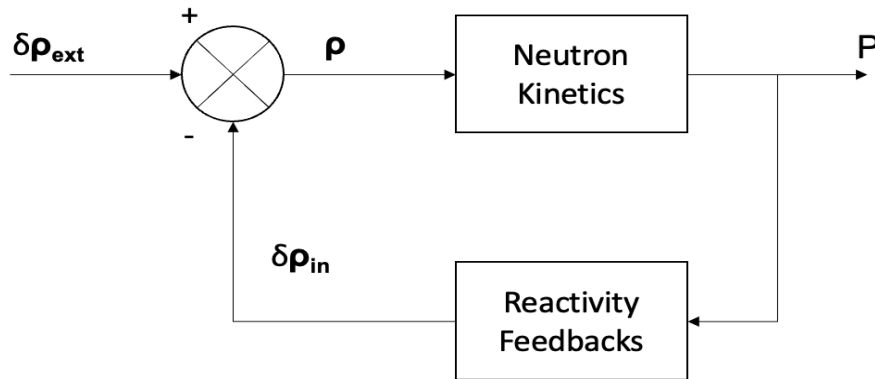
Let's start by characterizing the reactor power (P) as a closed-loop control system. We will decompose the reactivity (ρ) into two different reactivities: 'External' (ρ_{ext}) and 'Inherent' (ρ_{in}):

$$\rho_T = \rho_{\text{ext}} + \rho_{\text{in}}$$

Where

ρ_{ext} Reactivity change due to actions directly available for the operators in the control system (e.g. inserting/withdrawing control rods).

ρ_{in} Reactivity change due to the reactivity feedback effects.



The reactivity feedback mechanism will generate an effect in this closed loop, depending on how the specific parameter (e.g. temperature) of each major core component (e.g. moderator, fuel) influences reactivity. For instance: the moderator temperature reactivity effect.

Therefore, the reactivity feedback mechanism for a specific component (e.g. fuel) will be characterized by a series of reactivity coefficients (the change in reactivity per unit change in a parameter of the reactor), which correspond to the partial derivative, as reactivity depends on multiple parameters, of the core reactivity with respect to that specific parameter.

For instance, the temperature coefficient of reactivity (α_t) would be defined by the sum of the partial derivatives of the core reactivity (ρ) with respect to the temperature of the different (i) core components (moderator, fuel, etc.)

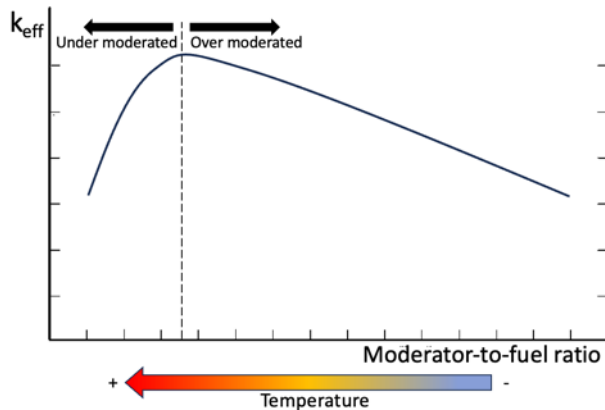
$$\alpha_t = \sum_i \frac{\partial \rho}{\partial T_i}$$

It is also worth noting that reactivity depends on multiple parameters simultaneously, so sometimes the reactivity feedback is better characterized by a coefficient of reactivity which is, in turn, the combination of different reactivity coefficients, such as the power coefficient.

There are three main reactivity coefficients that inherently contribute to the self-stabilization of a PWR, and therefore its safety:

1. Moderator Temperature Coefficient (MTC)

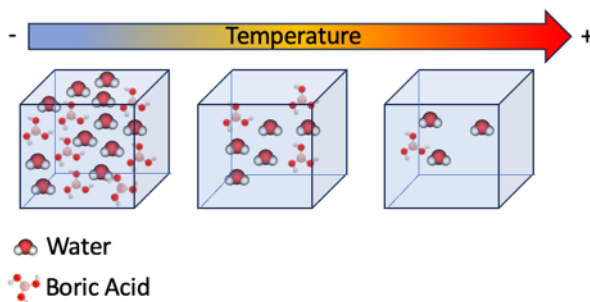
PWRs are under-moderated by design, so any variation in the moderator tends to work against neutron multiplication. In this case, when the temperature of the moderator increases, its density decreases (moderator-to-fuel ratio decreases, see diagram below).



Therefore, the neutron flux will find fewer opportunities to be moderated, and negative reactivity will be added.

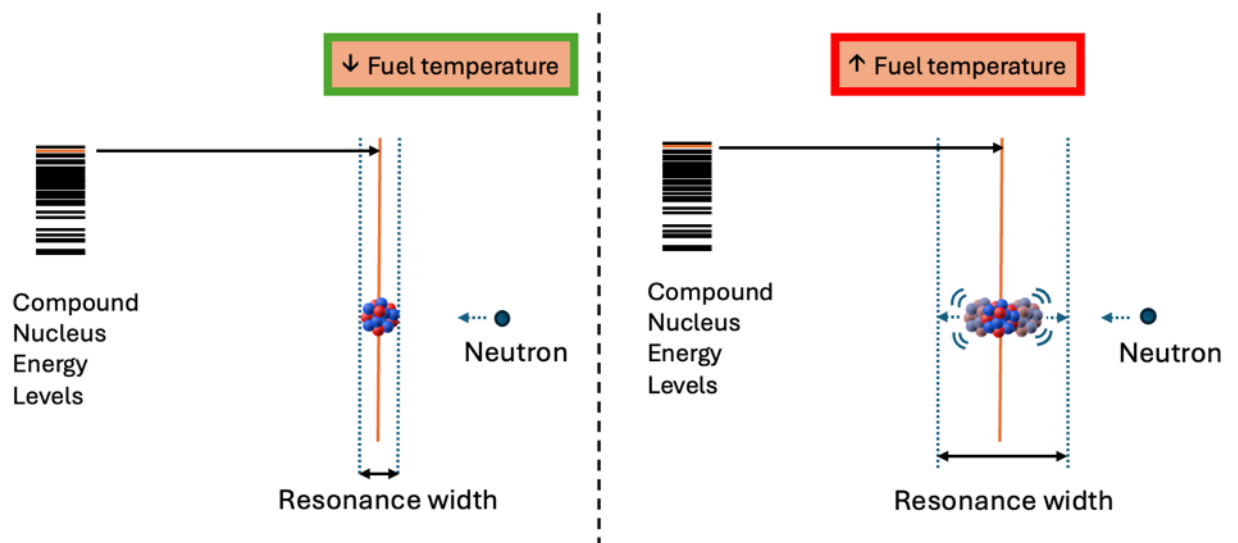
However, there is a specific situation where positive reactivity may be added instead: when the moderator contains a significant amount of soluble neutron poisons such as Boron.

As the diagram on the right shows, when the temperature increases the boric acid will be 'removed', thus adding positive reactivity. We will see a practical example of this scenario in the section on cyber-physical attacks.



2. Fuel Temperature Coefficient (FTC)

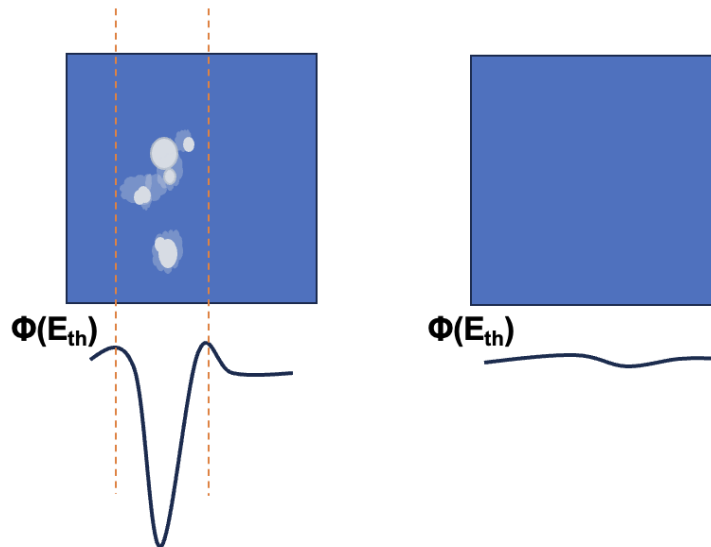
The most important phenomenon behind the fuel temperature coefficient is the Doppler-Broadening of resonances. A PWR designed with a negative FTC will prevent power excursions when the temperature in the fuel increases, because the doppler broadening effect will increase the probability of absorbing neutrons, thus rapidly decreasing the neutron flux.



3. Void Coefficient (VC)

PWRs (under-moderated) are designed with a negative void coefficient. Under regular conditions, the moderator should contain a minimal amount of voids (steam bubbles), otherwise a series of safety issues may arise, as we will see later. On the other hand, it is worth noting that the RBMK reactors of Chernobyl had a positive void coefficient.

When voids form in the moderator, for instance due to a decrease in the pressure, its density is effectively decreased, so there will be fewer particles available to slow down fission-generated neutrons. Therefore, a negative void coefficient will tend to stabilize the reactor as the temperature of the moderator increases, or the pressure decreases, because negative reactivity will be added. We can visualize the idea in the following image, by comparing the neutron flux of thermalized energies in the moderator with and without voids.

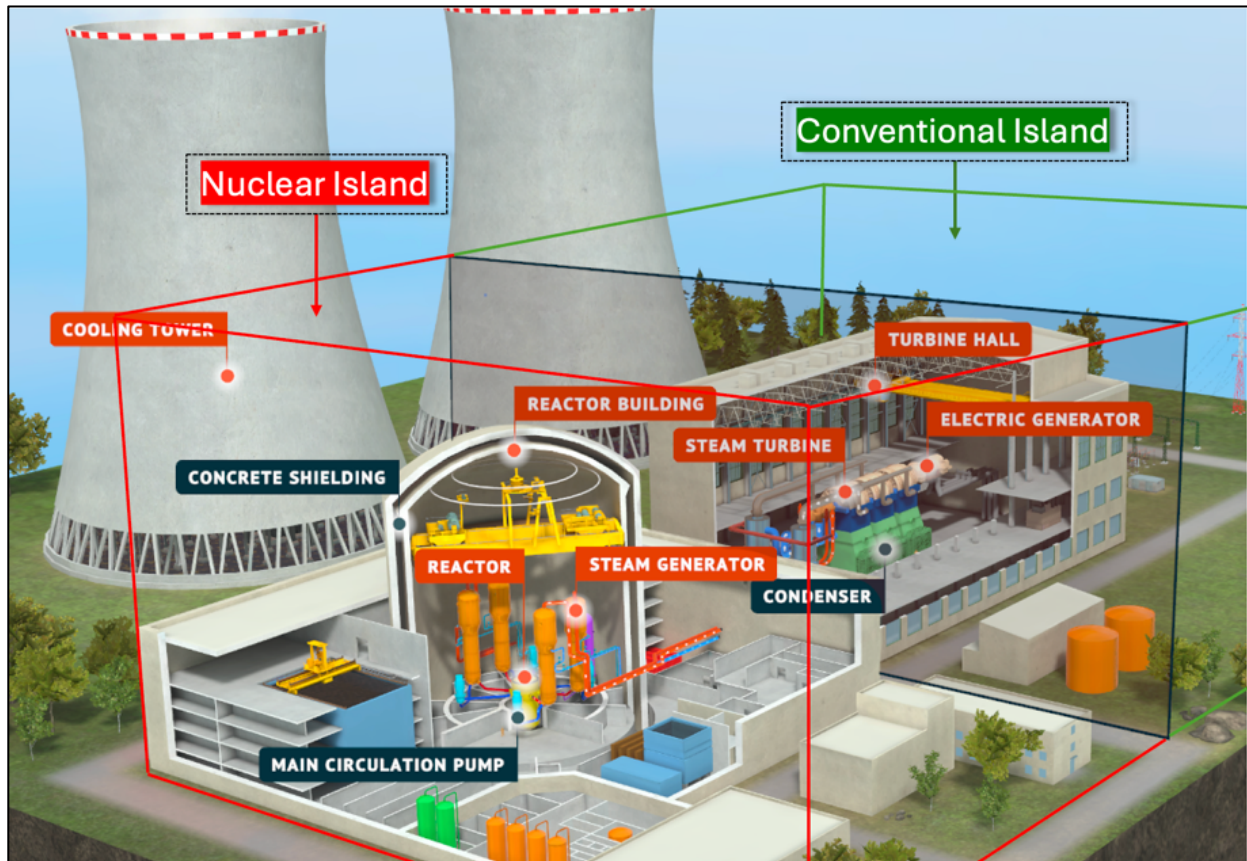


As we have just seen, temperature variations are the main factor that drive the most important reactivity coefficients, because of the induced changes in density (VC or MTC) or due to the Doppler broadening effect (FTC).

PWRs are inherently safe due to these coefficients, which are designed to rapidly, especially in the case of the FTC, compensate for any potential increase in reactivity.

1.4 Nuclear Power Plants

The main purpose of a nuclear power plant is to generate electricity, sharing the same operating principle as those thermal power plants that are also based on the Rankine cycle. The fundamental difference is how the heat is generated: thermal power plants burn fossil fuels as the heat source, while nuclear power plants use fission.



9

As a result, a nuclear power plant can be separated into two different parts:

1. Nuclear Island

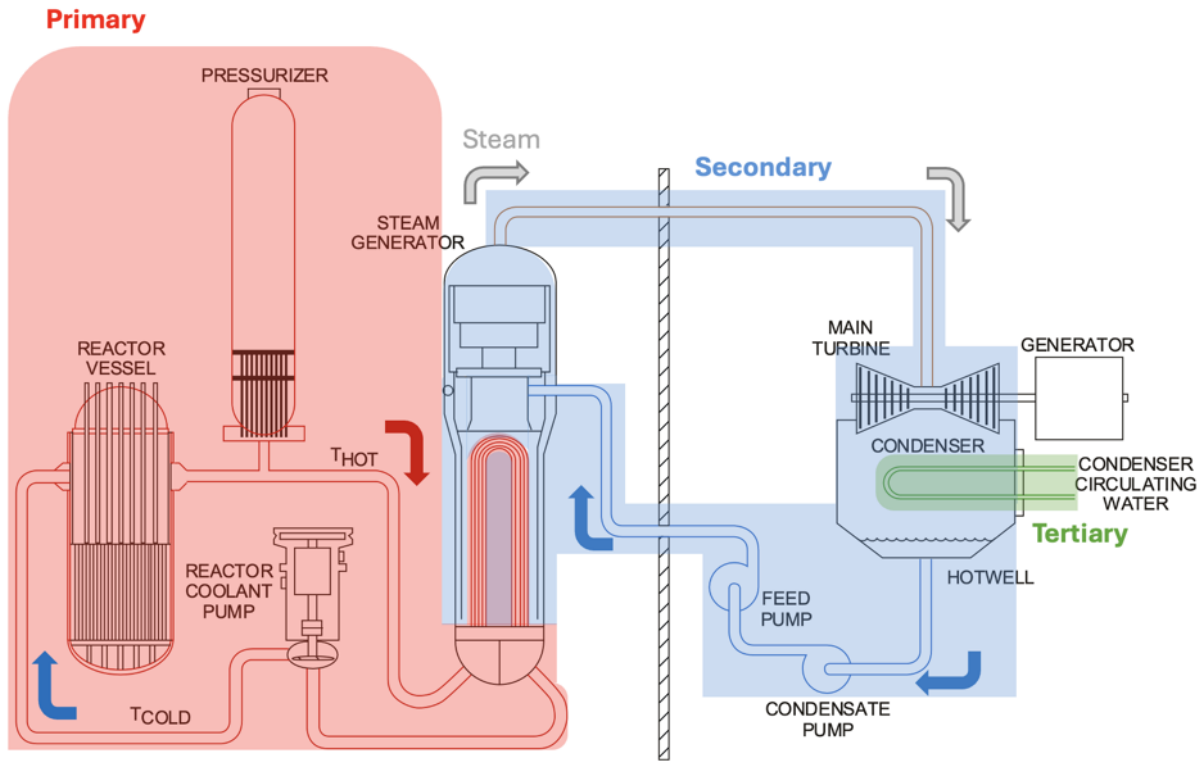
Here we find the Nuclear Steam Supply System. This consists of the nuclear reactor, most of its support, operation, control and safety systems, as well as all those components necessary to produce the steam that flows towards the turbine.

2. Conventional Island

This part would be similar to any other Rankine-based thermal power plant, although it may also contain systems and components that contribute to the safety of the reactor.

⁹ Certain elements of this image have been obtained from <https://www.energyencyclopedia.com/en/free-downloads/images>

A NPP with a PWR is comprised of three different circuits.



1. Primary (Closed loop)

This is also known as the 'Reactor Coolant System'. The coolant that is flowing inside the reactor vessel (basically demineralized water mixed with boric acid) removes the heat generated by the nuclear fuel in the core. This water is maintained at high pressure by the pressurizer to increase the boiling temperature, thus enabling, among other things, greater efficiency during the heat transfer process.

The heated water then flows through the Hot leg (T_{hot}) towards the steam generator's U-Tubes¹⁰, where its heat is transferred to the feedwater (secondary circuit). These U-Tubes act as a barrier between the radioactivity present in the primary coolant and the secondary circuit.

The now cooler water leaves the steam generator, being directed to the reactor coolant pump that increases the flow of water to keep the core properly cooled. Finally, the colder water returns to the reactor vessel through the cold leg (T_{cold}), thus completing the circuit. As a side note, the difference between T_{hot} and T_{cold} is probably less than one would expect, approximately 30°C.

¹⁰ For instance, the steam generators in Babcock&Wilcox's PWRs or VVERs, follow a different design, referred to as the 'Once Through Steam Generator'.

Please note that even without working Reactor Coolant Pumps (RCP), the difference in density between the hot leg and the cold leg enables natural circulation, but it's not capable of removing enough heat from the core at full power.

The primary circuit can consist of a different number of Steam Generator loops (hot leg, cold leg, Steam generator and RCP), depending on the reactor power. On the other hand, there is just one pressurizer that is connected to one of the hot legs through the surge line (which also accommodates potential water expansion if coolant temperature increases). Although not all elements of the primary loop have the same pressure, the changes in pressure generated by the pressurizer impact all of them equally as coolant flowing through both hot and cold legs mixes in the core.

2. Secondary (Closed loop)

In the steam generators the feedwater is in contact with the U-Tubes, thus enabling heat transfer between both circuits. The pressure of the water in the secondary circuit is lower than in the primary, so this heated water will boil, becoming saturated steam.

This steam then flows through the Main Steam Isolation Valves, which we can consider the limit between the nuclear and conventional islands, towards the turbine.

There are high and low-pressure turbines, which are moved by the different regimes the steam goes through. These turbines share a common shaft that drives the generator, producing the electricity that will be injected into the grid.

The condenser reuses the exhausted steam coming from the low-pressure turbine, by condensing it using the cold water coming from the tertiary circuit. The condensate pumps heat and filter this condensate and direct it towards the main feedwater pumps, where its pressure is increased, and it's also heated. Eventually, this feedwater returns to the nuclear island.

3. Tertiary (Open loop)

The coolant in this system is usually water collected from a nearby natural source such as a river, lake or the sea. The condenser circulating water is used to condense the steam, the corresponding acquired heat can then be removed in two ways:

1. Through cooling towers. What we see coming out of these massive cooling towers located at some nuclear power plants, is basically innocuous steam.
2. Water will be pumped back to its origin, without containing any kind of radioactive material, just a bit warmer.

For example, this has been used by different organizations to track the status of North Korea's experimental Light Water Reactor (LWR). By using satellite imagery of the Nyongbyon complex¹¹, it was possible to detect the discharge of warm water from the tertiary system, which would indicate that the reactor had reached criticality:

23. From mid-October 2023 until mid-March 2024, the Agency observed an almost continuous strong water outflow from the LWR's tertiary cooling water system. During a period of cold weather in mid-December 2023, ice melt in the river and steam from the water outflow were observed, indicating that warm water was being discharged and that the LWR had reached criticality.¹⁹ From mid-March 2024, the LWR was shut down for approximately 30 days, and since mid-April 2024, it has operated intermittently. These observations are consistent with the start of a commissioning process in October 2023, which has continued through the end of the reporting period.

(IAEA)¹²



(United Nations)¹³

In addition to these components, nuclear power plants have a series of support, auxiliary and emergency systems designed to ensure operational safety, regardless of the plant state.

¹¹ https://en.wikipedia.org/wiki/Nyongbyon_Nuclear_Scientific_Research_Center

¹² <https://www.iaea.org/sites/default/files/gc/gc68-15.pdf>

¹³ <https://documents.un.org/doc/undoc/gen/n24/032/68/pdf/n2403268.pdf>

1.4.1 - Safety

It is a fact that, for decades, fossil fuel power plants have been freely releasing tons of harmful emissions into the environment, thus greatly contributing to the extreme climate patterns we face today. These plants also generate a significant amount of the total fine particle air pollution, which is known to contribute to hundreds of thousands of premature deaths annually. This situation was downplayed for a long time, leading to the public perception of this continuous catastrophic anomaly as 'normal' or 'expected'.

By contrast, nuclear power plants are specifically designed to prevent any kind of radioactivity release or hazardous emissions into the environment. Unfortunately, as we all know, this is something that has not always been fulfilled, leading to shocking accidents on rare occasions, in very specific circumstances.

The three fundamental safety principles of a NPP are known as the three C's:

1. Controlling the reactor (reactivity control)
2. Cooling of the fuel
3. Confining the radioactive material

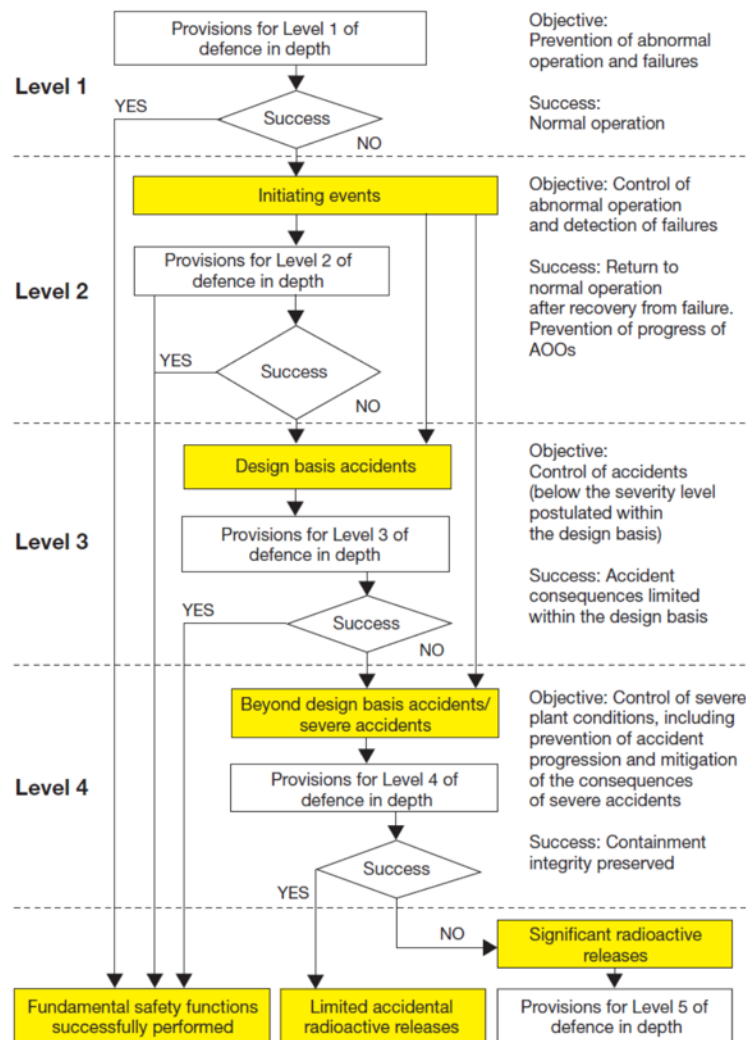
Defense in Depth (DiD) is a key concept for achieving these goals. The IAEA defines its implementation as follows:

Defence in depth is implemented primarily through the combination of a number of consecutive and independent levels of protection that would have to fail before harmful effects could be caused to people or to the environment. If one level of protection or barrier were to fail, the subsequent level or barrier would be available. When properly implemented, defence in depth ensures that no single technical, human or organizational failure could lead to harmful effects, and that the combinations of failures that could give rise to significant harmful effects are of very low probability. The independent effectiveness of the different levels of defence is a necessary element of defence in depth.

Therefore, DiD drives every single aspect of the design of a nuclear power plant, from physical, mechanical, and structural elements to fuel management, engineering, and Instrumentation and Control (I&C) systems.

The analysis of the following DiD flowchart¹⁴ provides a practical way to introduce some of the key safety concepts in nuclear power plants.

¹⁴ Assessment of Defence in Depth For Nuclear Power Plants
<https://www.iaea.org/publications/7099/assessment-of-defence-in-depth-for-nuclear-power-plants>



At least 5 different levels can be identified. Please note that although the basic principles are shared among countries, different regulatory entities may approach a DiD implementation in different ways, for instance subdividing levels or adding additional ones.

The transition from one level to another, which would correspond to the different plant states, requires a failure in the provisions¹⁵ for that specific level.

¹⁵ Measures implemented in design and operation such as inherent plant characteristics, safety margins, system design features and operational measures contributing to the performance of the safety functions aimed at preventing the mechanisms from occurring.

Level 1: Normal operation

This operational state is what is expected under any of the normal conditions for the plant, including shutdowns. Therefore, the objective at this level is the prevention of abnormal conditions and/or system failures.

If something went wrong with the provisions, then we would transition to the next level in the DiD hierarchy.

Level 2: Anticipated operational occurrences

Provisions in this level are intended to detect failures and control abnormal conditions caused by an 'initiating event'. This is an event that causes a disturbance in the normal operation with the potential to lead to core damage.

Therefore, at this level, the objective is to prevent the ongoing Anticipated Operational Occurrence (AOO) from progressing. An AOO is the deviation of an operational process from normal operation. It is expected to occur at least once during the operating lifetime of a plant but, in view of appropriate design provisions, does not cause any significant damage to items important to safety or lead to accident conditions.

Please note that, as introduced in the outline of Level 1 of DiD, shutdowns are considered part of normal operation. Therefore, a reactor scram (trip) can be the ideal way to handle an AOO at this level.

This is the last operational state, once the transition to level 3 occurs because of a failure at this level, we would be talking about accident conditions.

Level 3: Design basis accidents

At this point, things can go badly wrong. The objective at this level is either to prevent damage to the reactor core, and the release of radioactive material (which may require off-site protective actions) or minimize the consequences of such an accident as well as attempting to prevent its potential escalation.

Therefore, in the current state, the plant is facing a design basis accident: a postulated malfunction leading to accident conditions for which a facility is designed in accordance with established design criteria and conservative methodology, during which releases of radioactive material are kept within acceptable limits.

This is where all kinds of safety systems come into play, including the plant's inherent safety features as well as the I&C safety systems, such as the Reactor Protection System and the Engineered Safety Features Actuation System.

Level 4: Severe accidents and Design Extension conditions

Now the plant is facing a complex situation where the objective is to deal with the ongoing accident in order to both mitigate its consequences and limit its progression.

This is a critical point because the characteristics of the accident involve a serious degradation of the reactor core, including worst-case scenarios such as a core melt. As a result, structural and physical barriers are one of the pivotal provisions to implement this stage in the DiD

Level 5: Post-severe accident situations

Fukushima and Chernobyl are the only nuclear accidents that ended up with a large release of radioactivity into the environment. As a result, this kind of situation is no longer limited to the perimeter of the nuclear power plant but requires off-site mitigation plans and strategies.

Therefore, the purpose of this final level of defense is to act to reduce the radiological consequences of a large release of radioactive material or an early release of radioactive material that could potentially result from an accident.

1.4.2.- Instrumentation and Control

There are two types of safety systems: active and passive.

Active systems need to be powered by an external power source. In contrast, passive systems use natural forces or physical phenomena such as gravity (e.g. control rods), convection (heat removal) or pressure (e.g. pilot-operated relief valves). As a result, passive systems will perform their action inherently, in any given case.

On the active side, I&C systems play a fundamental role in DiD implementation, as they are designed to provide for the safe, secure, reliable and deterministic behavior of the nuclear power plant, especially in the first four DiD levels. This means that a NPP is greatly dependent on I&C systems for monitoring, supervision, control and protection.

I&C systems affect every aspect of the plant's operation, regardless of its state, but not all of them have the same impact with regards to safety. Therefore, the different regulatory entities and organizations make a clear distinction between safety and non-safety I&C systems (with an optional number of additional divisions in between, such as systems important to safety or safety related systems). The highest level of this classification is the "safety" I&C system (1E).

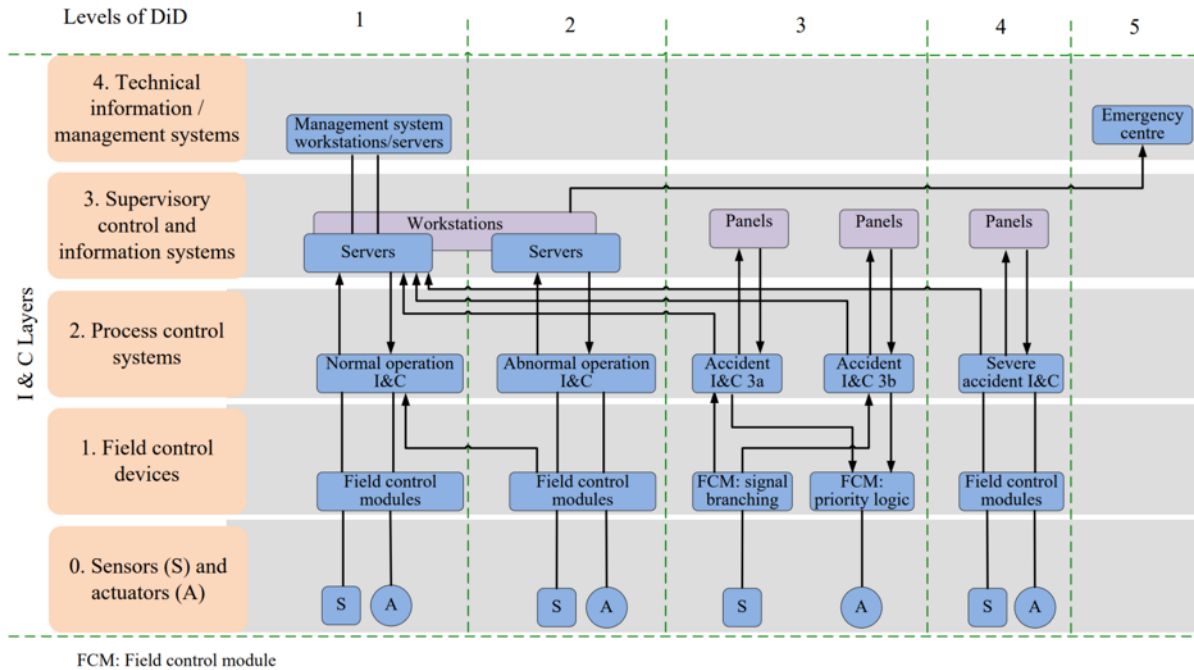
National or international standard	Classifications
IAEA	Safety system Safety related system Systems not important to safety
IEC	Category A Category B Category C Unclassified
France N4	1E 2E Important for safety (unclassified)
EUR	F1A (Automatic) F1B (Automatic and Manual) F2 Unclassified
Russia	Class 1 (Beyond DBA) Class 2 (Safety system, DBA) Class 3 Class 4
UK	Category 1 Category 2 Unclassified
USA (IEEE)	1E Non-nuclear safety

Safety systems are subject to the most stringent controls, in terms of hardware, software, redundancy, diversity, reliability and independence. Depending on the regulations and the plant design, there are usually certain restrictions (e.g. one-way communication, point-to-point data connections, deterministic traffic, and so on) applied for those networks containing a mix of these systems and others having a lower classification.

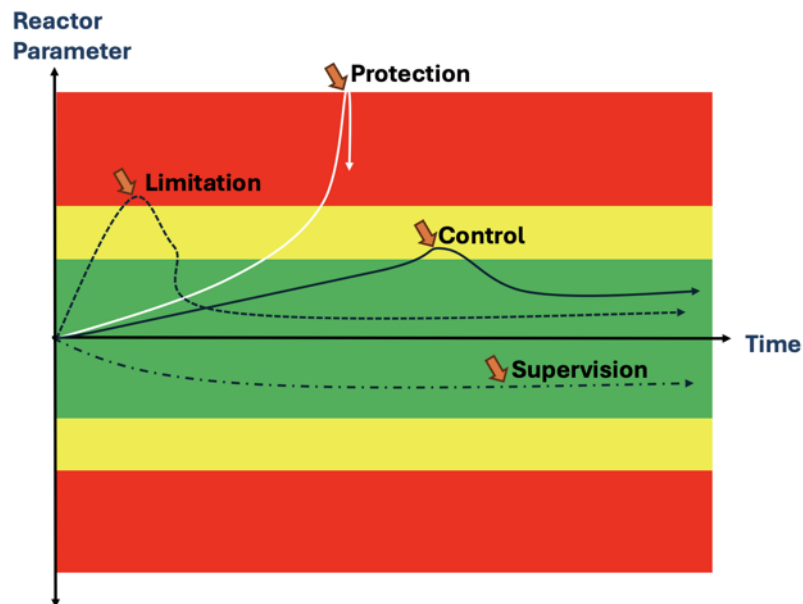
We should bear in mind that there is an important, intrinsic, relationship between safety and non-safety systems: a non-safety system cannot impact the safety of the plant, but it can act as a precursor, causing an initiating event that will eventually require the activation of a safety system.

In the table below (IAEA¹⁶) we can see a clear picture of the overall functionality. On the left column we find the different I&C layers, which are then mapped into the levels of DiD previously laid out.

¹⁶ https://www-pub.iaea.org/MTCD/publications/PDF/PUB1821_web.pdf



We can think of this process as a series of key reactor parameters that are being monitored over time, triggering both automatic and manual actuations when their values begin to fluctuate beyond their expected limits.



Level 1

The plant is operating at normal conditions. The plant display and monitoring systems will be providing the operators with continuous information about the different plant variables and key parameters, thus allowing them to supervise the conditions of the current operational state. If a

process parameter starts to follow an abnormal trend, the operators may be alerted by these systems to perform some control action.

Level 2

The plant has, somehow, entered into an abnormal condition. Therefore, it's time for the plant control systems to carry out a series of limitation actions (Reactor Control and Limitation System) that can bring the plant back to a normal operational state.

Level 3

In general terms, non-safety systems have been dominating the previous levels. This is the moment when safety systems need to react and perform their protective functions.

The Reactor Protection Systems (RPS) and Engineered Safety Features Actuation Systems (ESFAS) come into play, serving as the main line of defense against AOOs that could escalate into a severe accident. The scope of the research presented here is focused on these two safety systems.

At this level, the RPS will most probably trip the reactor and/or activate certain sequences of the ESFAS.

Level 4

The plant is no longer operating in a safe state. Reaching this level means a failure in the digital I&C safety systems. All the potential resources are then utilized in an attempt to prevent, or at least, mitigate a core melt and its corresponding release of radioactive materials.

The plant operators will now require a continuous flow of information to know the state of certain crucial safety parameters of the plant, in order to properly evaluate the situation. These variables are displayed on different safety Video Display Units.

The Post-Accident Monitoring (PAM) system, a safety system, provides this functionality and as such is also included (TXS QDS) in the scope of this research.

Level 5

The I&C systems are largely out of the scope at this point, as the worst-case scenario has transpired. Ideally, it is important to keep emergency communications and monitorization over the plant to the maximum extent possible from an alternative local, or even remote, control room or emergency center.

2. Actors and motivations

Article 56 of the Geneva Convention reads as follows:

1. Works or installations containing dangerous forces, namely dams, dykes and nuclear electrical generating stations, shall not be made the object of attack, even where these objects are military objectives, if such attack may cause the release of dangerous forces and consequent severe losses among the civilian population.

But also contains certain exceptions,

The special protection against attack provided by paragraph 1 shall cease:

[...]

b) for a nuclear electrical generating station only if it provides electric power in regular, significant and direct support of military operations and if such attack is the only feasible way to terminate such support.

As a result, the 'legality' (assuming an armed conflict) of attacking an NPP will always be subject to interpretation. Hence, it would be too optimistic, even unrealistic, to think that any international legal convention can provide any deterrence with regards to attacking nuclear reactors, either by cyber or kinetic operations.

This common dichotomy of employing cyber or kinetic operations against specific targets may have a different perspective in case of nuclear reactors, as their physical barriers are specifically designed to withstand powerful kinetic attacks (e.g. a commercial or military aircraft crash). As a result, it is reasonable to assume that this would be one of the few cases where it might be more feasible, assuming the attackers have capabilities to carry out an attack using both kinetic and cyber means, to cause physical damage by using a cyber approach, rather than a kinetic one.

Due to this complex context, the safety systems of nuclear reactors are not the kind of targets we usually see being attacked. In fact, there are no public reports of any cyber-physical attack of this type to date. This can be seen as a positive trend that might be used to, optimistically, conclude that we should not expect an attack against these systems. It is true that it would come as a surprise to suddenly see a wave of operations targeting NPPs, as opportunistic cyber-attacks can be pretty much disregarded.

Any serious attempt to successfully target a nuclear reactor to cause physical damage, by using cyber capabilities, will require a significant amount of resources and intelligence. A scenario usually limited to state-sponsored actors.

The motivation to undertake such a massive effort needs to be proportional to the ability of its authors to deal with, or assume, the implications of this first-of-its-kind operation.

And so, if that time ever comes, who is really capable of performing this kind of operation? I think that we can all acknowledge that there are nation-state actors with the technical capabilities and resources to carry out such an endeavor. As a constant in this research, I will endeavor to avoid plain opinions as much as possible, choosing educated guesses instead, so I will attempt to sustain this claim with some verifiable data.

By studying Stuxnet and Trisis, two of the most well-known ‘cyber-physical’ operations (although coming from totally different actors), it is possible to see how some of the patterns that emerge are similar to what would be required to carry out cyber-physical attacks against nuclear reactors.

2.1 Stuxnet

The ‘417’ payload, the first approach used to attack Natanz is truly fascinating. It clearly demonstrates the level of sophistication, information known about the target, and resources made available for the authors of Stuxnet.

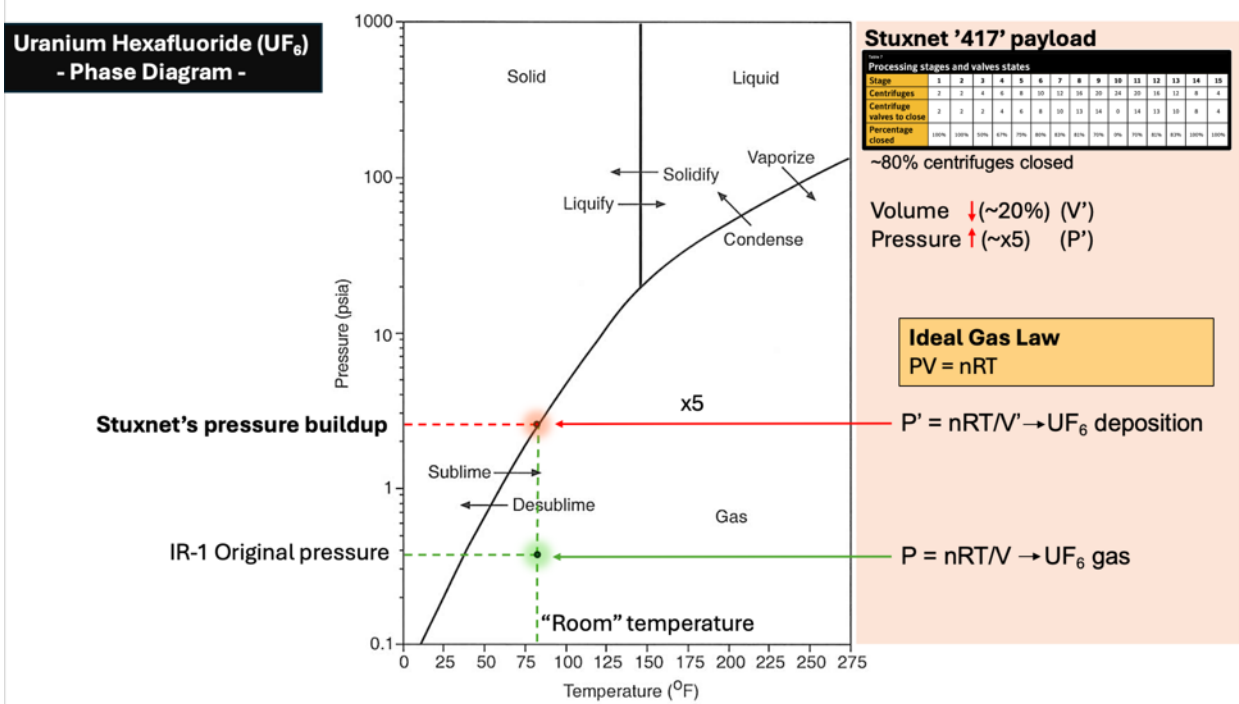
Langner¹⁷, Symantec¹⁸ and ISIS¹⁹ are the main points of reference to understand, what is considered, the first approach for attacking Natanz. The ‘417’ payload (which refers to the targeted Siemens PLCs) is surely the most sophisticated cyber-physical attack publicly documented. Let’s go over it briefly from a physics perspective.

In the image below we can see the phase diagram of Uranium Hexafluoride (UF_6), a key compound in the uranium enrichment process. At room temperature ($\sim 25^\circ\text{C}$), a moderate increment in the pressure would cause a change in the phase from ‘Gas’ to ‘Solid’, thus creating deposits of solidified UF_6 . This property was, allegedly, leveraged by the 417 attack to destabilize the centrifuge’s rotors. These would eventually ‘break’, earlier than expected, due to the imbalance generated by the centrifugal force derived from spinning a solid material instead of the expected gas molecules. Just imagine putting a brick into a washing machine.

¹⁷ <https://www.langner.com/wp-content/uploads/2017/03/to-kill-a-centrifuge.pdf>

¹⁸ <https://docs.broadcom.com/doc/stuxnet-missing-link-13-en>

¹⁹ <https://isis-online.org/isis-reports/detail/did-stuxnet-take-out-1000-centrifuges-at-the-natanz-enrichment-plant/>



However, to make this happen, first of all Stuxnet had to come up with a method to stealthily increase the pressure in the target centrifuges, without triggering the 'Cascade Protection System' (comparable to a 'Reactor Protection System' but several orders of magnitude simpler). This is a safety system designed to prevent and/or mitigate exactly that kind of scenario. We can easily make sense of Stuxnet's approach bearing in mind the Ideal Gas Law²⁰ and the following table included in the Symantec report.

Table 7

Processing stages and valves states

Stage	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Centrifuges	2	2	4	6	8	10	12	16	20	24	20	16	12	8	4
Centrifuge valves to close	2	2	2	4	6	8	10	13	14	0	14	13	10	8	4
Percentage closed	100%	100%	50%	67%	75%	80%	83%	81%	70%	0%	70%	81%	83%	100%	100%

For this purpose, let's provide some context:

- UF₆ is close to behaving as an ideal gas.
- The Feed (from which the UF₆ flow is injected) stage (a series of centrifuges connected in parallel) is the number 10.

²⁰ <https://www.britannica.com/science/ideal-gas-law>

Now, if we study the Table above, we can see how the Product (enriched UF_6) and Tails (depleted gas) stages (1 and 2, and, 14 and 15 respectively) are completely closed, meaning that the feed of UF_6 cannot 'escape' the cascade. Therefore, the pressure will gradually increase in the different stages of the cascade (both from left to right) towards the Feed Stage, as the feed flow rate is kept constant.

Another thing we can conclude from the table, is that approximately 80% of the total centrifuges will be isolated ('Centrifuge valves to close' in the table) during the attack. From the Ideal Gas Law we know that the Pressure is inversely proportional to the volume so we have:^{* 21}

- Before the attack: $P = nRT/V$ where V can be approximated, for simplicity, as the number of centrifuges times its volume
- After the attack: $P' = nRT/(V')$ where V' is approximately 20% of the total volume as close to 80% of the centrifuges have been isolated.

So according to the Ideal Gas Law, we have it that the resulting pressure in the rearranged cascade should be ~5 times more than the expected pressure, which is precisely the setpoint used by Stuxnet, and the pressure at which the UF_6 will start its desublimation.

State 4:

State 4 waits for the desired pressure change or other predetermined time limits before proceeding to state 5. If any of the following conditions are met, the code will continue to state 5:

- The pressure of the stage 10 or stage 11 transducer (these are likely transducers for or near the feed stage) has an absolute value greater than 280 units above the expected value and greater than five times the expected value.

To achieve this, Stuxnet implemented a very precise sequence that, among other things, involved opening certain relief pressure valves, usually controlled by the 'Cascade Protection System', to set back the pressure of the targeted cascade stages to regular levels.

2.2 Trisis

In contrast to Stuxnet, Trisis was discovered in-the-wild without a sophisticated cyber-physical payload. It did, however, provide a striking example of the successful targeting of digital safety systems in the industrial world for the first time. One of the main questions, that still remains publicly unanswered, is whether the petrochemical plant, which went into fail-safe mode after an unsuccessful and 'noisy' Trisis infection stage, was just a testbed or the real target. Bearing in mind that the Saudi plant was not the only victim of Trisis, it would be valuable to know the other sectors (if any) where Trisis was detected.

²¹ *For simplicity we obviate the volume of piping elements.

From the perspective of this research one of them is outstanding: Trisis targeted Triconex-based safety systems, which are, in addition to Teleperm XS, one of the few digital Instrumentation & Control platforms certified by the US NRC (and other nuclear regulators) to provide safety systems for nuclear reactors.

Some of the main features implemented in the Trisis implant can be easily extrapolated to Teleperm XS (TXS) due to a number of security-related characteristics that were similarly implemented in both Triconex-nuclear and TXS. Let's look at a relevant example of one of these functionalities that was effectively targeted by Trisis.

1. Firmware Integrity checks relying on CRCs

In both platforms, there is a lack of cryptographically secure mechanisms to validate the firmware.

TRICONEX - Topical Report ²²

Compiled application programs are downloaded to the Tricon V10 through a communication module. Programs and translated code are protected by 32-bit CRC. During the download process, the individual communication blocks have CRC protection. Communication blocks with computed CRC that does not match the transmitted CRC are rejected. In addition, the program segments, which may span communication blocks, have an overall 32-bit CRC. The 32-bit CRC for each program is stored both in the TriStation and in the Tricon V10.

The user may request a comparison between the content of the Tricon V10 and the data stored in the TriStation to be confident that the application in the Tricon V10 and the application last downloaded through the TriStation are identical. Comparison failures would indicate that the application in the Tricon V10 and the content of the TriStation are no longer the same.

Teleperm XS - U.S EPR Design Certification²³

Method of Verification 2:

Self-test qualification and configuration control: The TXS system software, including the software used in the self-test process, is developed and tested using a quality program as described in Reference 10. This verifies that the self-test features function properly. TXS system software contains an identification file providing a CRC [cyclic redundancy check] checksum for all files which are delivered within a package (e.g., executable programs, dynamic-link libraries, object modules, pre-links, header files). The CRC checksum of the complete TXS system software installation forms a unique identification of the version. When the TXS system software is loaded onto the TXS processing unit, the CRC checksum of the loaded TXS system software on the TXS processing unit is manually verified to match the CRC checksum of the originally developed and tested TXS system software. This verifies that the system software containing self-test features is identical to that which was tested and verified to operate correctly.

Based on these findings, it seems reasonable to assume that the same actors behind Trisis would be able to adapt it to target nuclear reactors, if in fact they haven't done so already.

²² <https://www.nrc.gov/docs/ML1209/ML120900890.pdf>

²³ <https://www.nrc.gov/docs/ML0907/ML090780501.pdf>

2.3 Targets

Attacking the safety systems of a nuclear reactor is a destructive operation per se, and if we ever see this kind of attack - or even just an attempt at it - in the wild, a red line will have been crossed forever. Paradoxically, despite all kinds of cautions, there is probably no completely 'safe' way to attack safety systems. There are so many things that can go wrong once you start messing with the last digital line of defense of a nuclear reactor that the worst-case scenario should always be assumed. Despite the inherent risks, we can identify some logic, and benefits, behind this aim for destruction.

a) Non-proliferation

Once again, Stuxnet is an obvious reference. Destroying specific reactors (e.g. clandestine experimental reactors able to breed weapons-grade plutonium), may provide a valuable advantage for certain actors looking to inflict damage on the nuclear weapons program of an adversary.

b) Military target

As we have seen in the Geneva convention, in the context of an armed conflict, nuclear reactors may be considered a legitimate target. This scenario may be especially important in the near future, when SMRs may be providing the energy requirements for specific military facilities. In this context, we should also consider other targets such as nuclear submarines.

Ongoing armed conflicts may bring some other complex scenarios. For instance, once an NPP that has been taken by foreign occupation forces is returned to its legitimate operators (i.e. Zaporizhzhya NPP), it shouldn't be disregarded that the digital Instrumentation and Control systems may have been backdoored.

c) Civil target

A successful attack against an NPP, or even a research reactor, would have a significant impact, not only in physical terms, but also from a psychological perspective. The way this can be leveraged by the attackers will depend on the specific context of the underlying conflict: demoralization, to spread panic or distrust in the government's capabilities, (pre/post)-war scenarios, economic damage, long-term sabotage during high demand seasons (winter, summer).

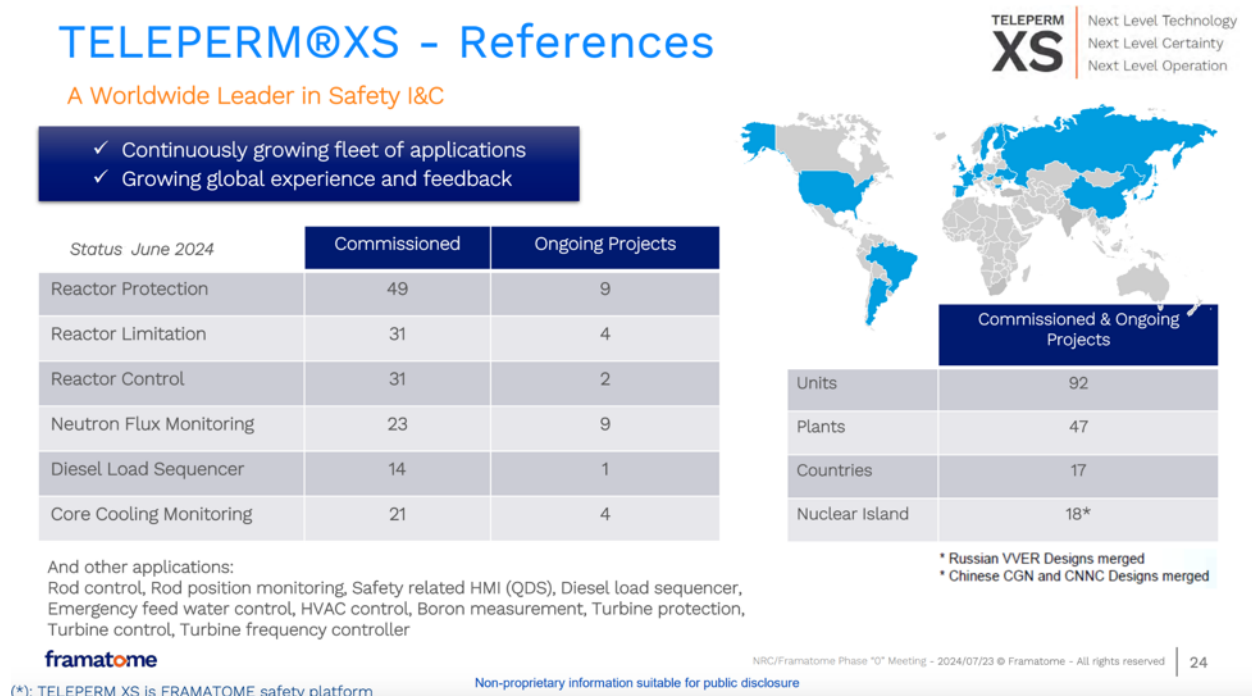
Let's not forget that AI's energy demands will soon be driving some of the global nuclear developments, both in traditional NPPs (e.g. Three Mile Island²⁴) and SMRs.

²⁴ <https://www.reuters.com/markets/deals/constellation-inks-power-supply-deal-with-microsoft-2024-09-20/>

3. Teleperm XS

Framatome's Teleperm XS (TXS) is a digital²⁵ I&C platform designed specifically for use in safety systems in NPP²⁶, as a replacement for, or upgrades to their analog counterparts.

It is one of the most widely used digital safety I&C platforms, sustaining the main defense line (RPS, ESFAS, etc.) in dozens of nuclear reactors globally, including Europe, USA, Russia, and China.



The TXS platform consists of four main elements:

- **Hardware**
There are dozens of available components, including digital/analog signal acquisition and conditioning, I/O, priority modules, actuation and processing computers, neutron flux instrumentation, communication modules, video display units, and so on.
- **System Software**
In the lower layer we have the operating system, which was originally developed by Siemens, and on top of this the platform software, where we find the runtime environment.

²⁵ <https://www.youtube.com/watch?v=WGjG1VRXk2k>

²⁶ <https://www.nrc.gov/docs/ML2419/ML24193A231.pdf>

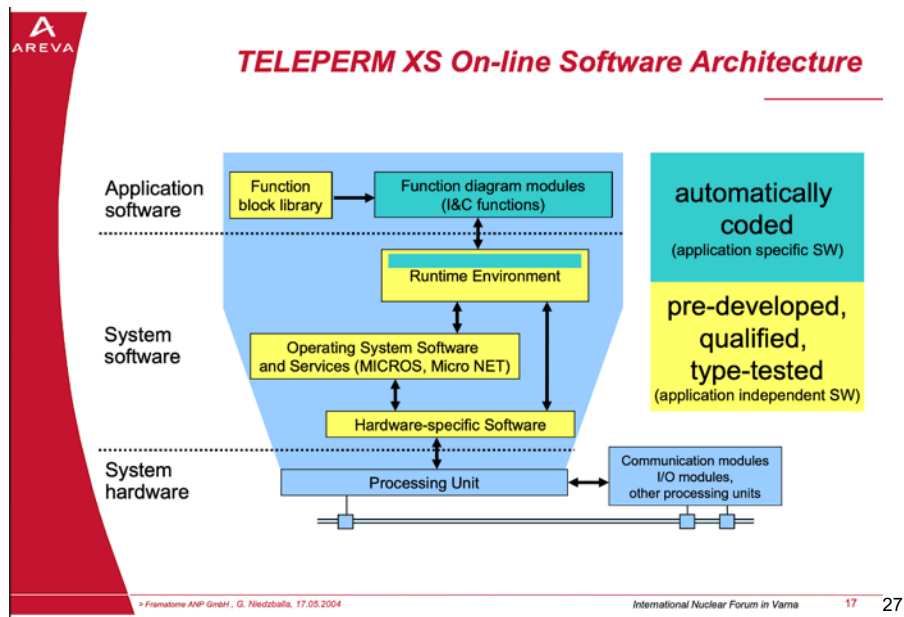
This latter component provides a unified environment for execution of the application programs.

- Application Software

The application software performs the plant-specific TXS safety-related functions using function block modules which are grouped into function block diagram (FBD) modules. Ultimately these FBD programs are translated into C and compiled to generate the resulting binary that will be loaded into the function processor.

- Engineering tools

The application software is generated (automatic code generation, FBD → C → x86) by a custom engineering and design tool (SPACE) which uses common, and plant-specific, function blocks to construct specific applications.



3.1 Public availability of Teleperm XS

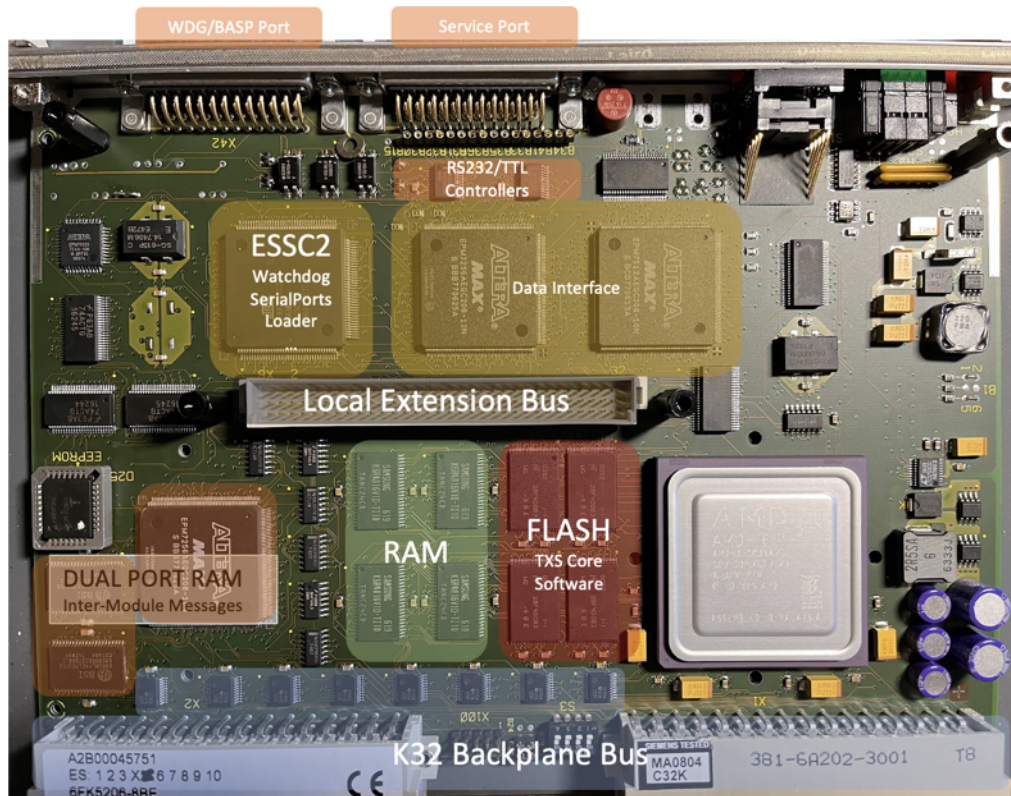
In September 2023, I noticed that a plethora of brand-new Teleperm XS (2nd generation) components were available on Ebay, marketed by different German and Dutch vendors. This public availability contrasted with the lack of components for other similar safety I&C platforms, also used in the nuclear industry. Therefore, it cannot be completely ruled out that the nuclear phase-out carried out in Germany may have plausibly resulted in a dump of stock to 'less restricted' sales channels.

²⁷ <https://www.nrc.gov/docs/ML0037/ML003711856.pdf>

Obviously, that was a good opportunity to dig deeper into the, usually, closed world of nuclear digital safety I&C systems, so I acquired the following Teleperm XS modules:

1. ‘SVE2’ – Main Processor Module

The main processing module (See Appendix A - Figure 2) for operation on the K32 TXS backplane bus. It is a microprocessor (AMD K6-2) based device, designed specifically for TXS.



2. ‘SCP3’ – H1 Communication module

This is the TXS-specific Communication processor (See Appendix A - Figure 1) for (H1) Ethernet. It consists of a SVE2 processing module, although loaded with a specific software, and the H1 communication module (image below) connected through the local extension bus.

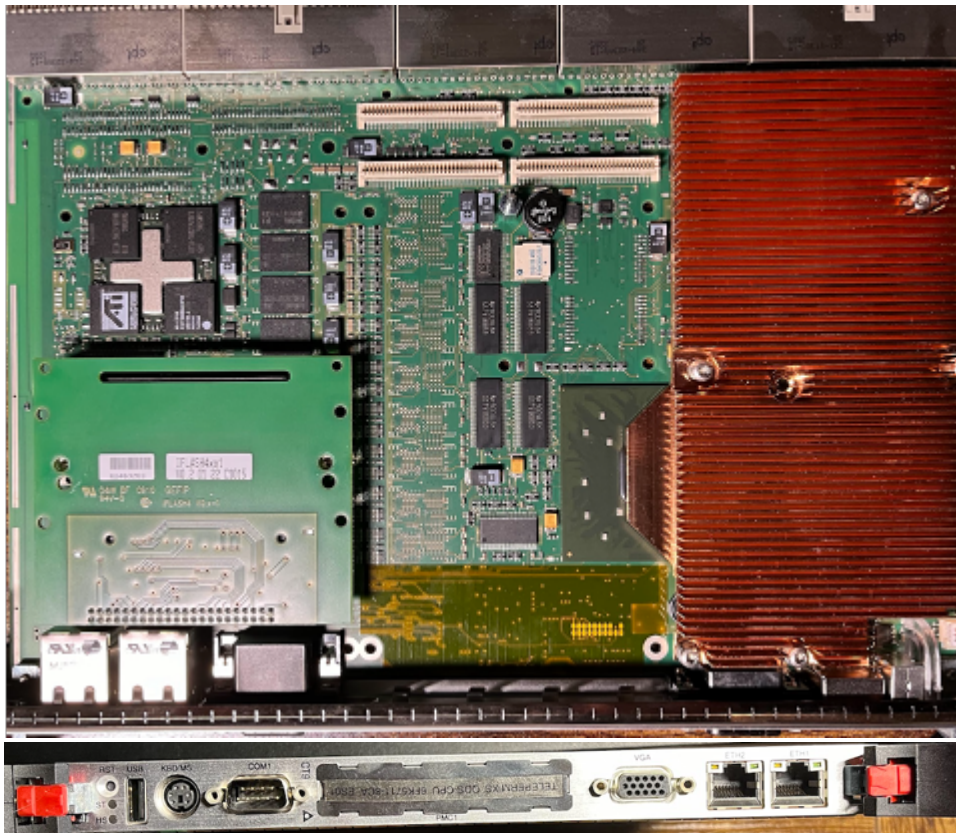


3. **‘SES1’ – Ethernet Switch and media converter**

A COTS Switch (Hirschman RS2-3TX/2FX²⁸) and media converter (optical-electrical for the required safety isolation) for TXS Ethernet networks.

4. **‘QDS-CPU’ – Qualified Display System CPU**

As opposed to the SVE2/SCP3, and in the line of the SES1, the QDS-CPU is another COTS device, a CR9 v3.x CompactPCI²⁹ single board computer manufactured by SBS Technologies.



The CompactFlash, easily accessible through the IDE board (iFlash4xx1), is expected to contain the OASIS-based QDS³⁰ system, when deployed in the field.

²⁸ https://www.doc.hirschmann.com/pdf/Anl_RS2xTXxFX_EEC_11_0415_en.pdf

²⁹ https://www.artisan-g.com/info/SBS_CR9_CP9_CT9_Manual.pdf

³⁰ https://www.researchgate.net/profile/Stephane-Louise/publication/228882206_The_OASIS_based_qualified_display_system/links/0deec52a0af7aa80e4000000/The-OASIS-based-qualified-display-system.pdf



5. **'QDS-SR' – Qualified Display System SubRack**

A COTS subrack for the QDS-CPU.

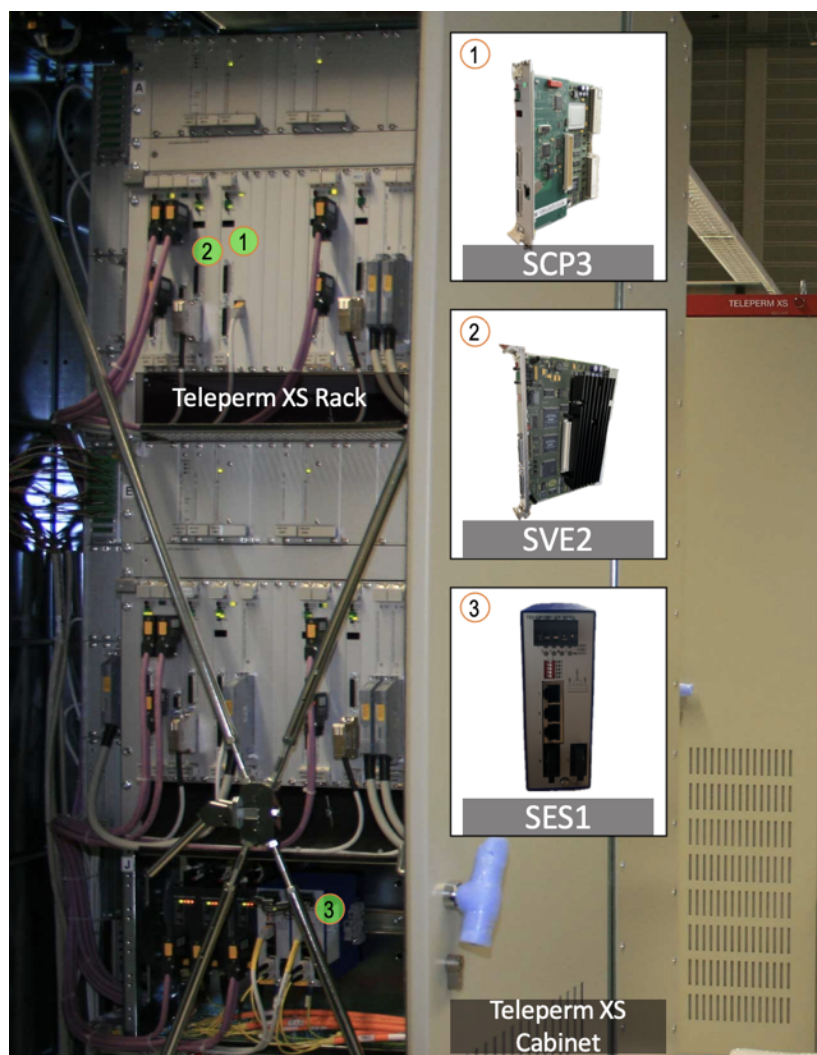
6. **'QDS-Display' – Qualified Display System Touchscreen**

A COTS Touchscreen that serves as a safety qualified Visual Display Unit. The QDS-Touchscreen contains a COTS Touchscreen controller (EGALAX ETP S4500), which transmits the operator interactions with the touchscreen to the QDS-CPU via an insecure serial protocol (RS232/COM1 port).



It is important to clarify that, unfortunately, these brand-new modules, 'SCP3', 'SVE2' and 'QDS-CPU', do not contain any kind of TXS-specific firmware. I dumped the flash and EEPROM memories from these modules to confirm it. The firmware loading phase for these TXS devices is carried out during the preparation for the Factory Acceptance Test (FAT), as we can read in the "Software Program Manual for TELEPERM XS"³¹.

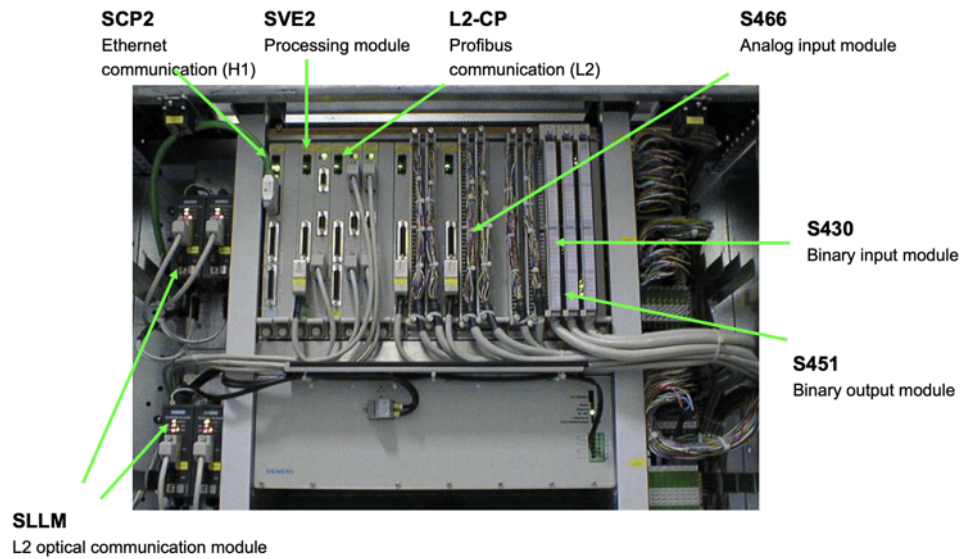
Software may be installed at several points during the course of the project. The first time an Application Software Release is installed is in preparation for the FAT. The initial software load is made using the TELEPERM XS Maintenance Laptop, since bootstrap loading of any TELEPERM XS processor is not possible via the TELEPERM XS Service Unit because access from Service Unit is not possible without TELEPERM XS System Software, Application Software, and pre-defined communication links Installed.



Teleperm XS 2nd Generation (new)

³¹ <https://www.nrc.gov/docs/ML1129/ML11294A397.pdf>

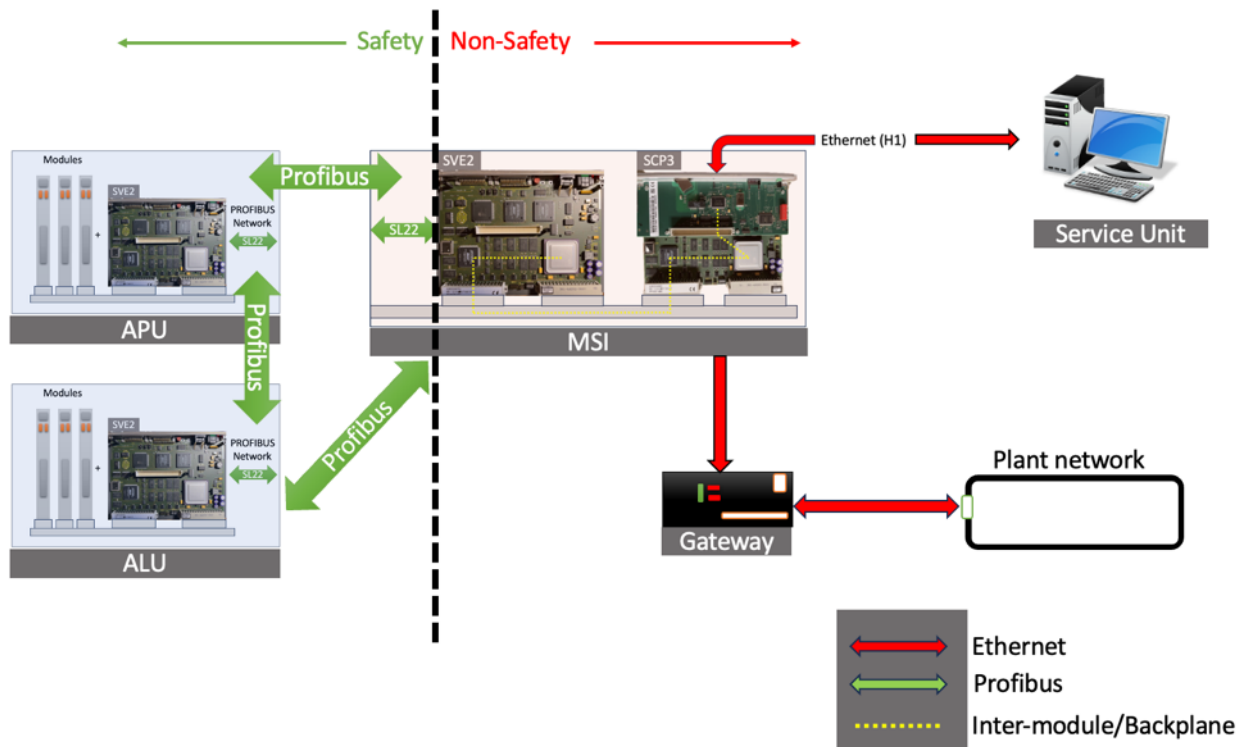
TELEPERM XS Hardware: Example Configuration



Teleperm 2nd Generation (legacy)

3.2 An overview of TXS

In order to approach the analysis of the attack surface in this complex platform, let's start by introducing the most important functional units that can be found in TXS-based architectures, for different NPPs.



1. APU (Acquisition and Processing Unit)

The APU consists of the SVE2, which acts as the function processor, different input and output modules, and a Profibus communication module.

APUs acquire the inputs from the instrumentation channels and proceed according to the implemented logic and configuration (e.g. setpoints). The result of these calculations is then sent to their corresponding ALUs.

2. ALU (Actuation Logic Unit)

In terms of hardware, the ALU is essentially similar to the APU, although its function is different.

ALUs perform the voting logic according to the redundant inputs received from their corresponding APUs. In safety critical systems, one of the fundamental requirements is redundancy to avoid spurious signals. The ALUs are in charge of this redundancy layer by implementing different voting rules and fail-safe conditions. When the ALU has completed the voting logic (e.g. 3 out of the 4 redundant APUs transmitted the same coherent output), it transmits the actuation orders to the corresponding field devices (through an AV42 priority module) and actuators.

3. MSI (Monitoring and Service Interface)

The MSI is the functional unit that acts as a communication node between safety and non-safety systems. On one side, it communicates with the Profibus network (safety) and on the other it receives data and commands from the non-safety, ethernet network.

Therefore, it has to act both as a barrier and a facilitator of the communications between the safety network and the 'outside world'. Without being completely accurate but making use of well-known terms, we could say that MSI serves as both a gateway and a firewall between the safety and non-safety networks.

From the cyber-security point of view, this is one of the key elements to take into account when exploring remote attacks.

4. Service Unit

The Service Unit (SU), a COTS computer running SUSE Linux, is one of the few units (if not the only one), that is granted bidirectional communication with the (Profibus) safety network from the (ethernet) non-safety network. This communication is not direct but goes through the MSI, which also filters unwanted traffic.

The SU is able to perform critical tasks on the previous units, such as

- System diagnosis.
- Monitoring the system functional status.
- Performing periodic tests of the system (e.g. triggering specific actuation sequences).
- Modifying the changeable software parameters (e.g. setpoints).
- Loading new software.

This computer contains the required tools, configurations, settings, databases and interfaces that allow an almost complete control over the digital I&C system.

As a result, it is a primary target for any malware-based attack.

5. Gateway

An industrial computer is used as the interface to the operational I&C, to transmit data coming from the safety network (through the MSI). The Gateway needs to translate from TXS-specific protocols into standard ones, such as OPC. In general terms, its communication is unidirectional from the MSI and bidirectional with the plant network.

A compromised gateway would grant the attackers the ability to either hide or modify specific information or values (alarms, status messages, etc.) that will be consumed by the plant operators.

The lack of firmware in the acquired TXS components limited the extent of this research. Therefore, the analysis is mainly based on the observed hardware capabilities as well as the official documentation available. Additionally, I'm only considering attack scenarios with the following three characteristics:

1. Remote

Malicious actions have their origin in the non-safety part of the ethernet network, being able to reach their targets via the same network.

2. Malware

Malicious actions are initiated by autonomous software-based implants.

3. Initial access is assumed

It is assumed that the actors will have enough resources to plausibly compromise equipment required to gain initial access. The methods required to accomplish this task may include 0-days and/or supply-chain attacks against COTS/open-source elements such as operating systems or software packages (operating systems, databases, etc.).

This doesn't mean that other scenarios are not possible, but these are three characteristics we have already seen in real-world cyber-physical operations.

3.3 Exploring the attack surface

As has been shown, the 'brain' of a TXS architecture is the SVE2, which can be found controlling the logic in the three main functional units: APU, ALU and MSI. This means that the ability to read and write arbitrary memory (e.g. modify setpoints), as well as to execute arbitrary code (deploy an implant) in a selected SVE2 would be the ideal mechanism for implementing subsequent cyber-physical payloads.

Therefore, as it has been anticipated in the previous section, the Service Unit would be the primary attack vector because it implements the functionality required to perform this kind of action and

most importantly, it's usually the only functional unit with bi-directional communication with the MSI.

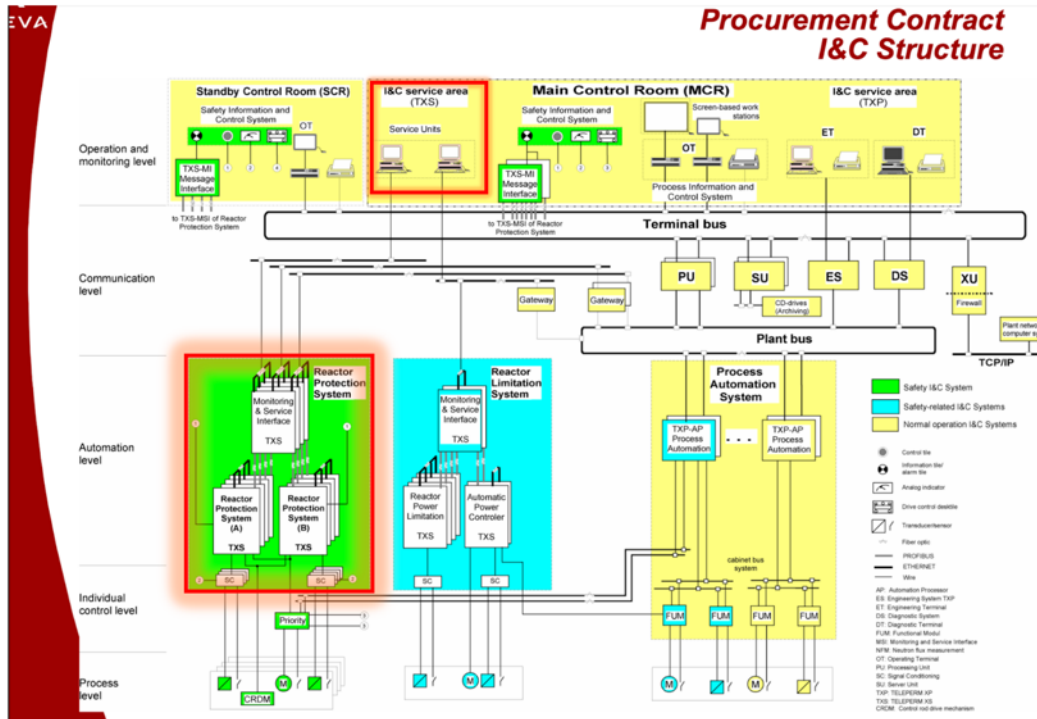
In order to provide a realistic scenario, I'll elaborate these objectives in the context of existing NPPs. This implies contrasting their architectures, design and mitigations with our objectives as potential attackers. It should be understood that this is just a theoretical exercise, mainly based on publicly available information. I'm not claiming, and therefore it cannot be claimed by using this research as a pretext, under any circumstance, that these nuclear power plants are vulnerable or insecure.

3.3.1 - Tianwan NPP and Russia's AES-2006

Let's start this journey in China. By 2027, Tianwan NPP is expected to become the largest nuclear power plant in the world, with up to 8 working units, mostly VVERs.

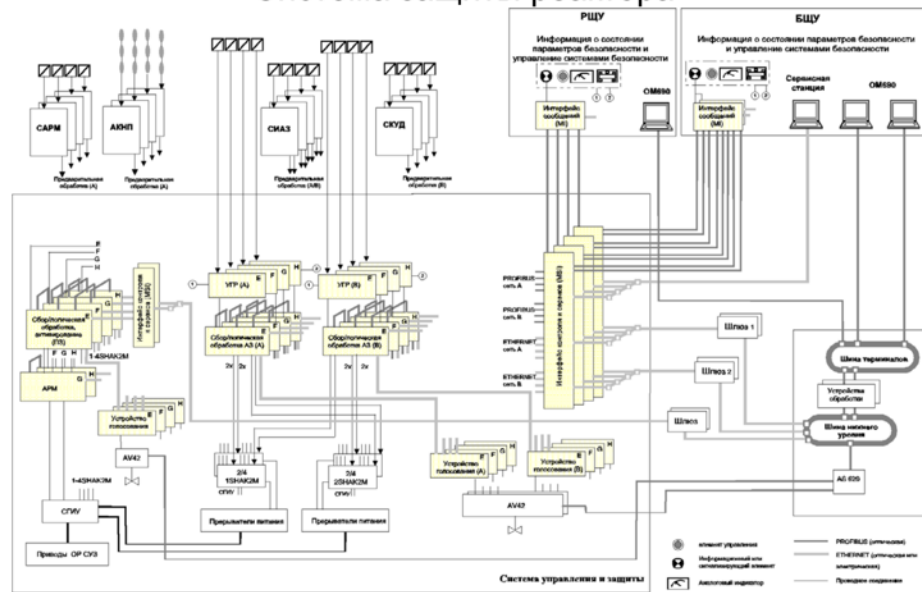


The TXS platform has been used to implement the RPS/ESFAS and the Reactor Control Limitation System. In the following image (slides from Areva), we can see the architecture of the I&C systems for Tianwan Units 1 and 2. I've highlighted two parts: "I&C Service Area" and the "Reactor Protection System", so let's analyze them.



TXS has also been a fundamental³² part of Russia's 'AES-2006'³³ design. Therefore, we can find a similar architecture in the RPS/ESFAS of Russian nuclear power plants such as Leningrad-2 or Novovoronezh-2.

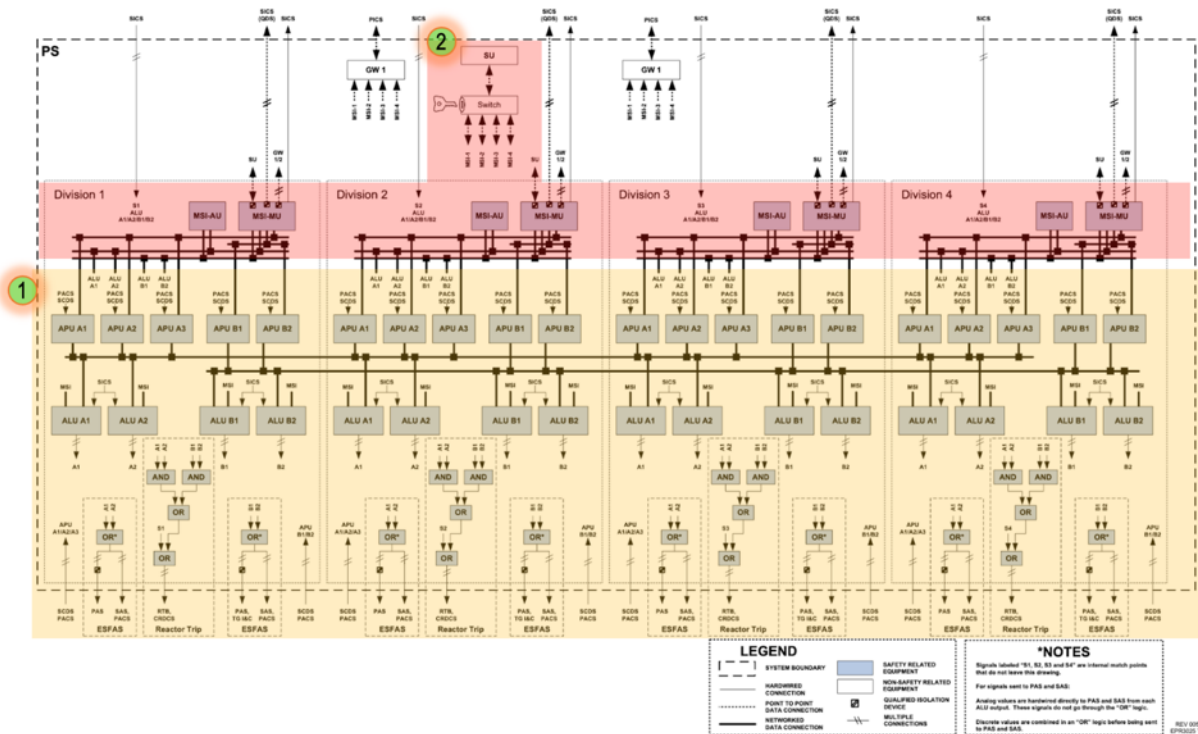
Система защиты реактора



³² <https://www.sa.aveva.com/news-russia-aveva-np-supplies-safety-instrumentation-and-control-system-for-generation-3-reactor>

³³ <https://ru.wikipedia.org/wiki/%D0%90%D0%AD%D0%A1-2006>

Basically Tianwan, and AES-2006 based NPPs, are equipped with the common 'Reactor Protection System' design that Areva implemented for its TXS platform. We can see a similar Areva design, but with a more detailed view in the following layout for the, now defunct, US EPR. Let's break it down in two parts.



1. The RPS is organized into four redundant divisions (which can house one or more channels*³⁴), physically located in different Safeguard Buildings (adjacent to the Reactor Building). Each division contains two functionally independent subsystems (A and B), which are redundant to the same subsystem in the other divisions (e.g. Subsystem A in Division 1 is redundant to Subsystem A in the rest of the divisions, so are their corresponding APUs/ALUs). Each subsystem may implement a different logic to protect against, usually, the same underlying condition by sensing different physical events (collecting measures from different sensors), thus also ensuring diversity at different levels.

The RPS A subsystem comprises 3 APUs (APU A1, APU A2, APU A3) and 2 redundant ALUs (ALU A1, ALU A2). The same applies for RPS B.

Therefore, the redundant ALUs in each subsystem process the outputs from their respective redundant APUs across divisions.

³⁴ An arrangement of sensors, data acquisition and signal processing electronics, I/O modules, logic units, and actuator or field devices that conform the automation path.

For initiating a “Reactor Trip”, the outputs of the redundant ALUs within the subsystem are combined using a hardwired AND. This prevents spurious signals from progressing. Then, the result of this AND for each subsystem is combined by using a hardwired OR (diversity) to generate a divisional Reactor Trip signal, which is then propagated to the reactor trip devices and control rod drive control system where they will actuate the corresponding element according to a predefined configuration.

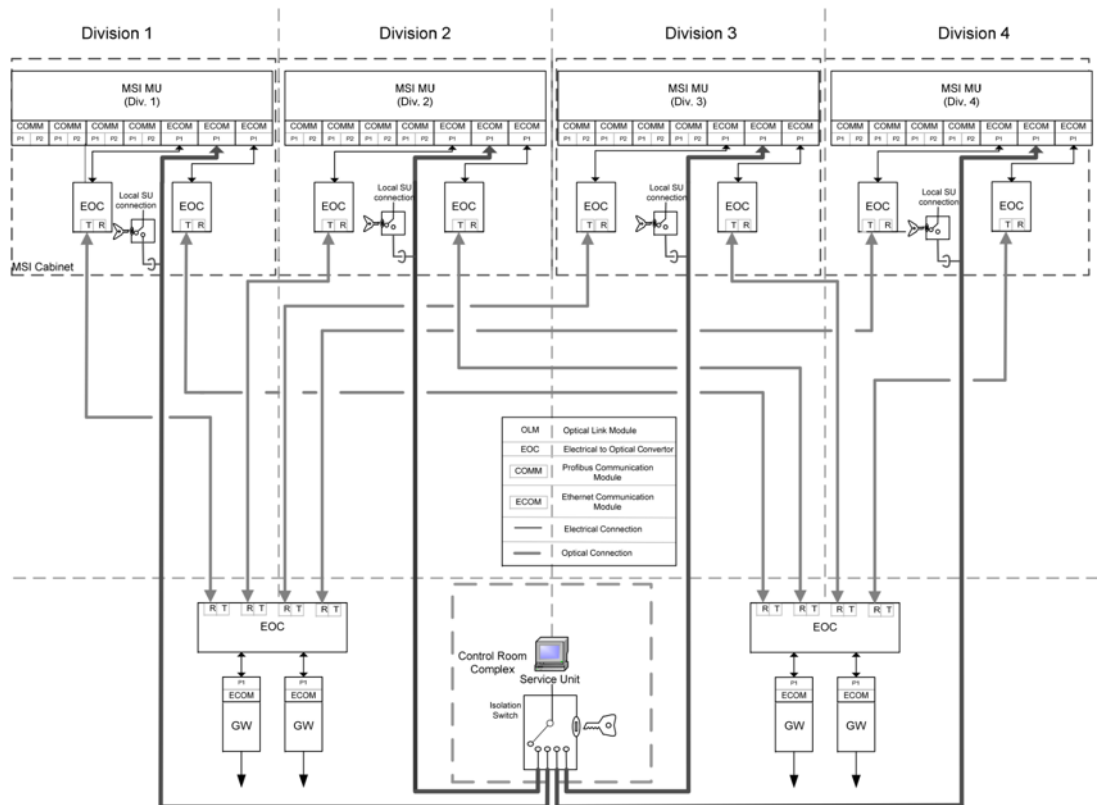
For initiating the “ESFAS” sequence, the output of each redundant ALU within the same subsystem is combined by using a hardwired OR. Thus, within a division each subsystem will be able to initiate an actuation order, which will be then propagated to the priority and actuation control system.

A series of additional hardwired ORs can be observed, those are included to reflect the possibility for operators to initiate manual trips and actuation sequences.

2. Each division has an MSI-MU (MSI Main Unit) that serves both subsystems and has a bi-directional ethernet connection (point-to-point in theory) with the Service Unit. This is the most important attack vector, as it could potentially enable malware coming from the non-safety network to reach the safety network. However, it wouldn’t be quite that easy.

In the diagram we can observe a key switch. It represents a physical key switch that isolates the communication path from the SU to just one division at a time (the one enabled by the key switch position). There is a central switch in the MCR and local switches in each division. It is important to note that, according to the US EPR FSAR³⁵, this switch would be hardwired, which means that it physically disconnects the communication flow. However, I could not find any positive indication that this had been implemented in the same way for Tianwan or AES-2006 NPPs. Actually, as we’ll see, in certain NPPs this physical disconnection does not seem to exist, relying on administrative controls instead.

³⁵ <https://www.nrc.gov/docs/ML1107/ML110760443.pdf>

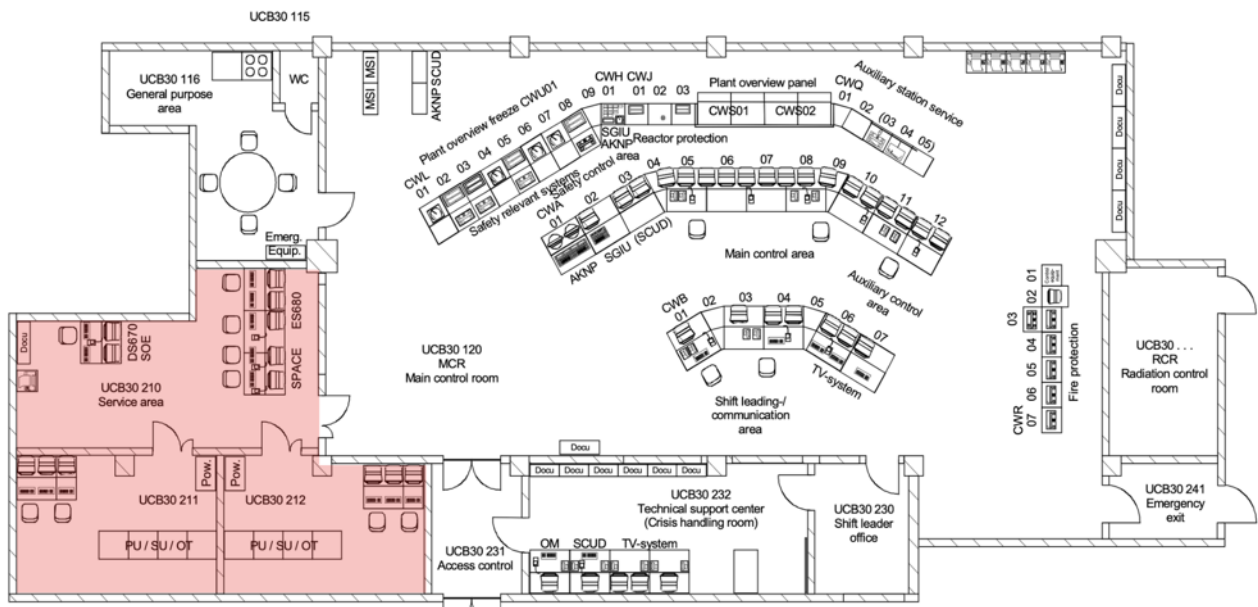


3.3.1.2 - I&C Service Center

The I&C Service Center is a separated area located within the Main Control Room (MCR). It contains the service unit (SU), as well as other engineering workstations with the tools (e.g. SPACE) and devices required by the technicians and engineers to perform maintenance operations, conduct periodic tests or upgrade the digital I&C systems.

It is worth mentioning that the personnel entering this room are expected to be different from the regular operators working on the MCR. However, they are still required to pass through the Access Control area.

The image below shows the architecture of the Tianwan NPP MCR. The I&C Service area has been highlighted.



Thanks to a video³⁶ that shows a tour on the Tianwan MCR, we can successfully match the architecture depicted in the plan, with its real-world implementation.



This I&C Service Center represents a common design in modern, highly integrated control rooms, so we can find it in many other NPPs. For instance, in Olkiluoto 3 NPP³⁷ (Finland).

³⁶ <https://youtu.be/mk1Tf2NkFlo?t=2508>

³⁷ <https://info.ornl.gov/sites/publications/files/Pub6813.pdf>

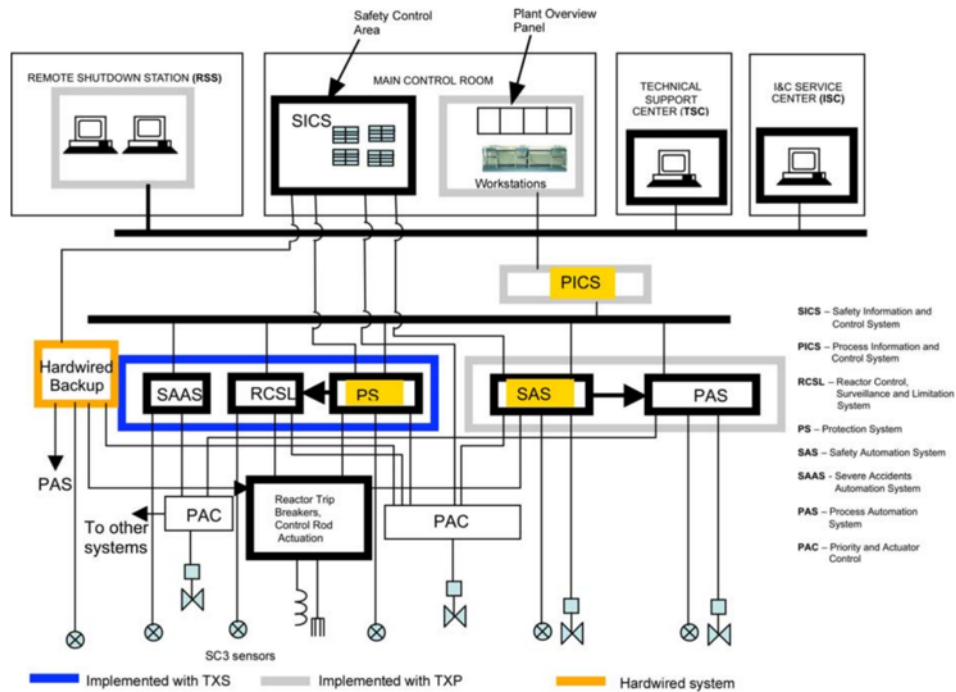
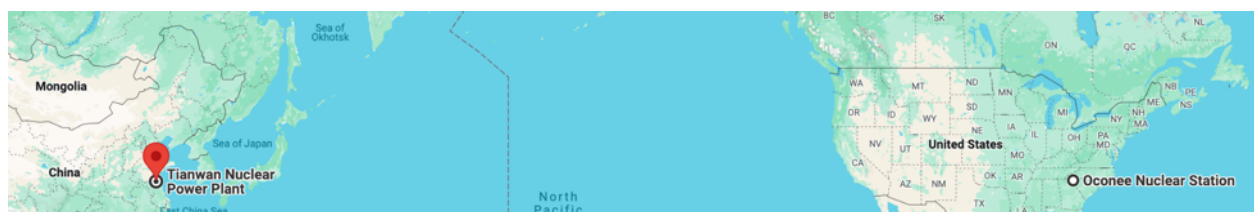


Fig. 3.10. Olkiluoto 3 I&C architecture. (Source: J. Hyvarinen, STUK)

Although there are different layers of access controls and mitigations, there is not anything yet that truly prevents a remote, malware-based operation. In view of the analyzed elements, we have it that:

- There is a bi-directional ethernet communication, ideally point-to-point and limited to a single division at a time, between safety and non-safety functional units through the MSI.
- The non-safety functional unit that enables this communication, the Service Unit, is usually located in a specific area of the Main Control Room. This area is designed for technicians and engineers, who can also be external contractors.

To continue our journey of exploration of the attack surface that would potentially enable this scenario we will virtually travel 11,500 km to a new stop: Oconee NPP.

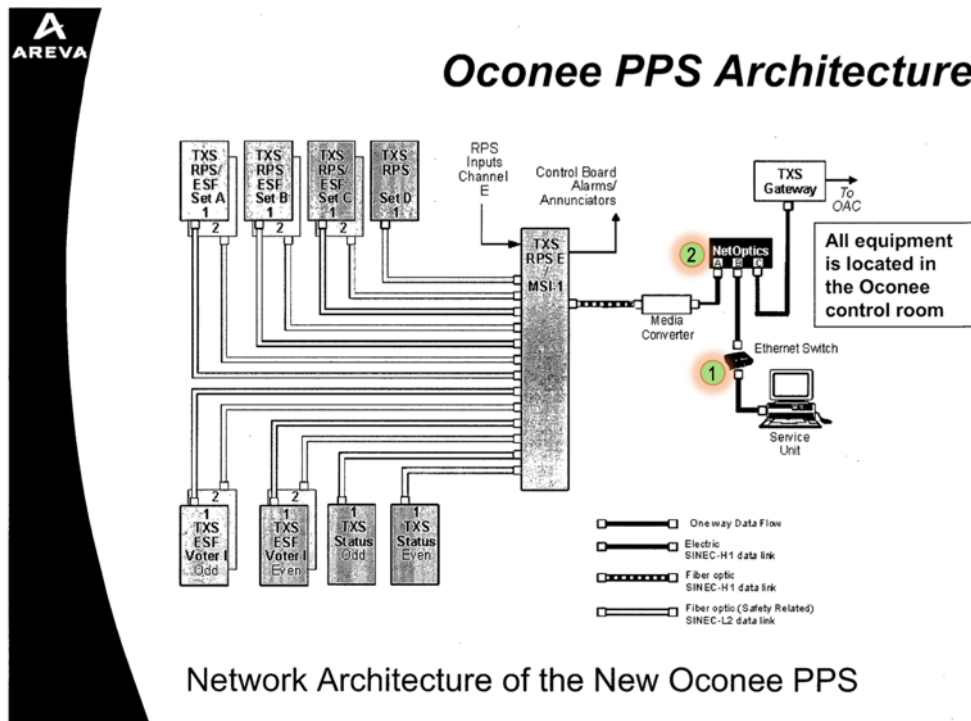


3.3.2 - Oconee NPP

In 2011, Oconee was the first US nuclear power plant to upgrade their analog I&C systems with a digital solution. As the reader would no doubt guess correctly, TXS was the platform chosen for this massive effort.

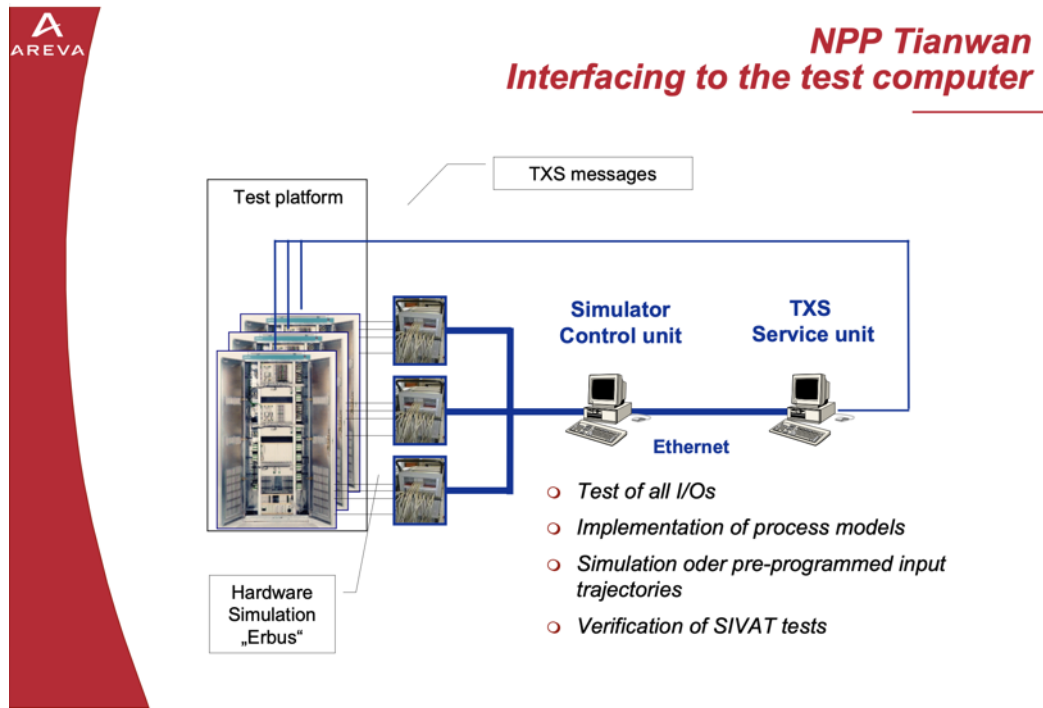


The architecture³⁸ of this digital Plant Protection System (RPS/ESFAS) presents some new elements we haven't seen previously.



³⁸ Oconee LAR <https://www.nrc.gov/docs/ML0807/ML080730339.pdf>

1. The communication path between the MSI and the Service Unit is no longer point-to-point. Instead, we can see an 'Ethernet switch', although I couldn't find any further information about what other devices are connected to it. However, based on other designs, it is likely that this switch is used to connect the Service Unit with the Simulator Control Unit and/or the SPACE workstation.



2. Additionally, a Netoptics Tap device is used to implement one-way links (e.g. MSI → Gateway), as well as to allow bi-directional ones (MSI ↔ Service Unit). We also observe a media (opto-electrical) converter, to provide electrical isolation for safety purposes.

It is interesting to note that, as reflected by the Oconee License Amended Request, the MSI implements an access control based on the MAC address of the Service Unit³⁹.

Communications on the non-safety related side of the MSI is through a restricted access local area network (LAN) that connects the MSI, TXS Service Unit computer, and the TXS Gateway computer. The MSI is designed and programmed to only relay control and maintenance commands that originate from the Ethernet media access control (MAC) address assigned to the TXS Service Unit.

This feature matches with what is described in the "Software Program Manual for TELEPERM XS Safety Systems Topical Report".⁴⁰

³⁹ <https://www.nrc.gov/docs/ML1013/ML101390464.pdf>

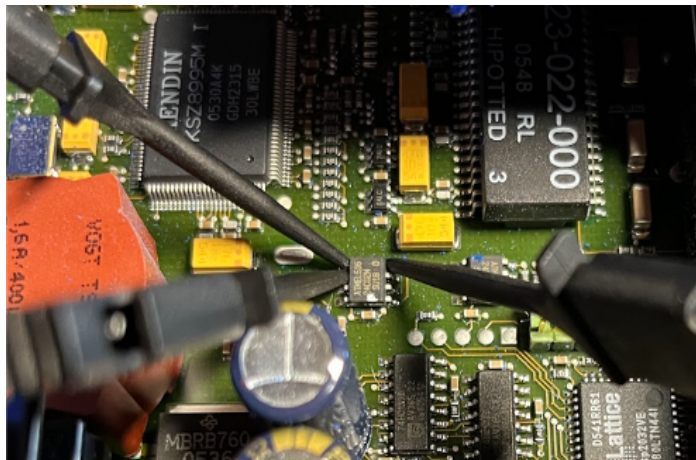
⁴⁰ <https://www.nrc.gov/docs/ML1129/ML11294A397.pdf>

C.4.2.2 TELEPERM XS Service Unit Access Control

The TELEPERM XS Service Unit is used by authorized personnel to monitor and test the TELEPERM XS System, to diagnose system alarms and failures, and to make parameter and software changes. The TELEPERM XS safety function computers are not directly accessed by the TELEPERM XS Service Unit; instead the TELEPERM XS safety function computers are only accessed via the MSI, which provides communication independence. The TELEPERM XS System and the associated TELEPERM XS Service Unit are located in secure areas. The identification of the TELEPERM XS Service Unit by the MSI is accomplished by a protected MAC Address used to decode incoming service messages.

In this context, it's relevant to provide some additional information about the TXS SES1 device, which was previously introduced. This device is a media converter and ethernet switch, based on a 'KENDIN KSZ8995M'⁴¹ integrated L2 switch component. The 'managed' part of this switch is implemented through an I2C EEPROM, from which the configuration for the SES1 is read. Besides this, the SES1 cannot be managed.

So, I tapped into the EEPROM while booting, thus capturing the configuration to confirm that it was being used. Essentially, most of the configuration registers contain the default values specified in the KSZ8995M datasheet⁴². This simple analysis revealed that there is not any kind of 'Port security' related configuration. Obviously, I also verified this feature by generating traffic with different spoofed MAC addresses.



However, it is important to clarify that I couldn't find any reference to confirm whether SES1 is used in Oconee.

⁴¹ <https://www.microchip.com/en-us/product/ksz8995>

⁴²

<https://ww1.microchip.com/downloads/aemDocuments/documents/OTH/ProductDocuments/DataSheets/ks8995m.pdf>

3.3.2.1 - The case of the “Staff Position 10”

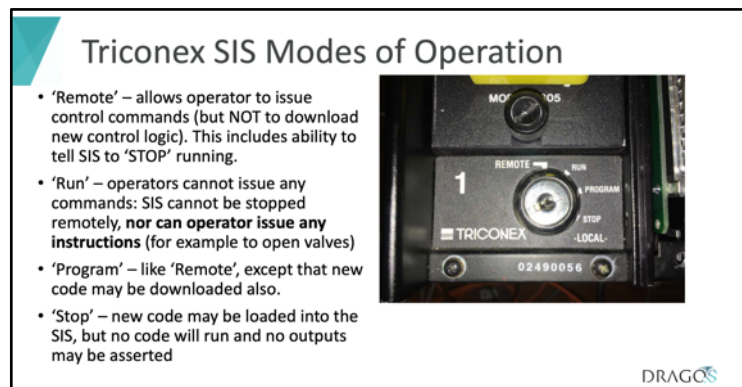
The US NRC Digital Interim Guidance⁴³ for ‘Highly Integrated Control Rooms & Digital Communication Systems’ established a series of requirements for licensees to get their digital I&C safety systems approved.

As one would expect, one of the main concerns is the bi-directional communication between safety and non-safety equipment. Those digital I&C platforms, such as Triconex or TXS, that were eventually approved, managed to ‘easily’ demonstrate conformance with most of the restrictions imposed by the US NRC, except for one: the Staff Position 10, which reads as follows:

10. Safety division software should be protected from alteration while the safety division is in operation. On-line changes to safety system software should be prevented by hardwired interlocks or by physical disconnection of maintenance and monitoring equipment. A workstation (e.g. engineer or programmer station) may alter addressable constants, setpoints, parameters, and other settings associated with a safety function only by way of the dual-processor / shared-memory scheme described in this guidance, or when the associated channel is inoperable. Such a workstation should be physically restricted from making changes in more than one division at a time. The restriction should be by means of physical cable disconnect, or by means of keylock switch that either physically opens the data transmission circuit or interrupts the connection by means of hardwired logic. “Hardwired logic” as used here refers to circuitry that physically interrupts the flow of information, such as an electronic AND gate circuit (that does not use software or firmware) with one input controlled by the hardware switch and the other connected to the information source: the information appears at the output of the gate only when the switch is in a position that applies a “TRUE” or “1” at the input to which it is connected. Provisions that rely on software to effect the disconnection are not acceptable. It is noted that software may be used in the safety system or in the workstation to accommodate the effects of the open circuit or for status logging or other purposes.

Here, the US NRC is describing a crucial defensive pattern to prevent unintended modifications of the safety logic, but also cyber-attacks. This is in line with the design requirements described by other organizations such as the IAEA⁴⁴.

PLCs, and industrial controllers in general terms, have different modes of operation, which delimit the features and logic available. The ‘front-end’ to change these operating modes is usually a mechanical element (e.g. a key switch), and although the US NRC is asking for a hardwired logic, the ‘backend’ is usually just software. In this context it’s unavoidable to mention an outstanding example of this scenario: Trisis⁴⁵.



⁴³ <https://www.nrc.gov/reading-rm/doc-collections/isg/digital-instrumentation-ctrl.html>

⁴⁴ https://www-pub.iaea.org/MTCD/publications/PDF/Pub1694_web.pdf

⁴⁵

https://recon.cx/2018/montreal/schedule/system/event_attachments/attachments/000/000/044/original/RECON-MTL-2018-DRAGOS_TRISIS_RECON2018.pdf

For some reason, in the Saudi petrochemical plant where Trisis was detected, the Triconex safety controllers were continuously operating in 'Program' mode. Leaving the controllers in this mode generates an alarm for the operators, but they were acknowledging it daily, probably without realizing its potential implications.

However, Trisis was also prepared for a less ideal scenario, as the implant was specifically designed to accept remote commands "regardless of the key switch position" (CISA report⁴⁶).

Implant

The implant, in many ways, is far more straightforward than the injector. This code is run when the compromised TS protocol command is received and provides RAT-like functionality. Most importantly, it allows an actor to read and write memory—including within the in-memory firmware region—and execute arbitrary code regardless of the key switch position, including "RUN." This allows an actor to effect changes on the controller while it is in full operation, not just while it is being reprogrammed. Figure 8 shows the control flow of this component.

The idea is pretty effective for attacking safety-related digital I&C systems: infect when possible, attack at any time.

This approach was possible because in the Triconex platform, the key switch does exactly the opposite of what the US NRC required: the logic to change operating modes is entirely handled by software without any hardwired logic, so a compromised firmware can just ignore the key switch position. This situation is explicitly noted in the final safety evaluation of the Triconex topical report⁴⁷.

However, the Tricon V10 keyswitch does not provide a physical disconnect or interruption of the connection by means of hardwired logic as required by ISG 4, Staff Position 1, Point 10, but instead sets 2 bits within the software to change the operating mode of the Tricon. Therefore, the Tricon V10 relies on software to effect the disconnection of the TriStation capability to modify the safety system software, a condition specifically stated as not acceptable in this ISG staff position. Based on the information provided in the LTR, the NRC staff determined that the Tricon V10 platform does not meet the NRC staff guidance provided in Staff Position 1, Point 10.

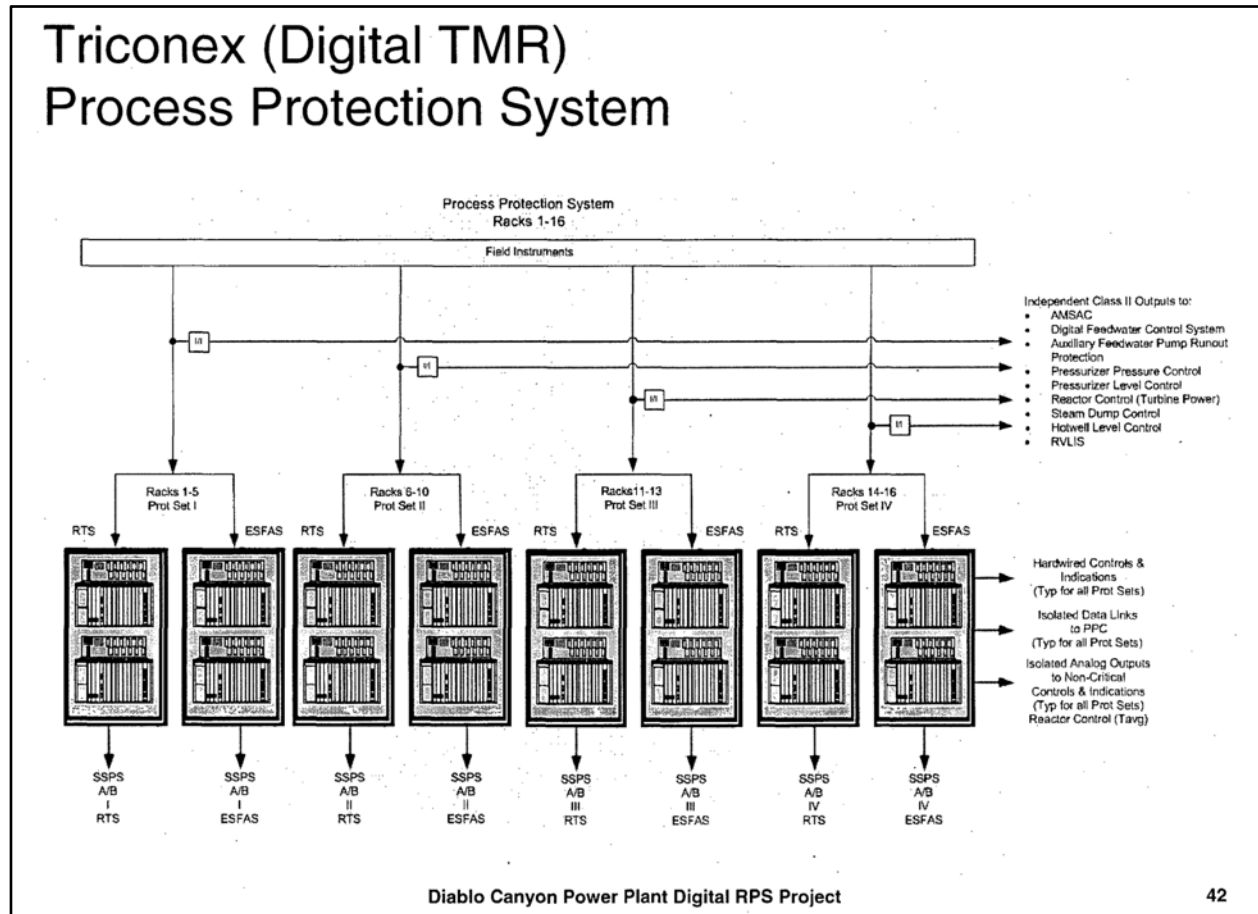
In order for the NRC staff to accept this keyswitch function as compliant with this Staff Position, the NRC staff will have to evaluate an application specific system communications control configuration—including the operation of the keyswitch, the software affected by the keyswitch, and any testing performed on failures of the hardware and software associated with the keyswitch. This is an ASAI associated with the implementation of this keyswitch.

⁴⁶ <https://www.cisa.gov/sites/default/files/documents/MAR-17-352-01%20HatMan%20-%20Safety%20System%20Targeted%20Malware%20%28Update%20B%29.pdf>

⁴⁷ <https://www.nrc.gov/docs/ML1209/ML120900890.pdf>

Ignoring the key switch position is precisely what enabled Trisis' backdoored packet handler to process the attacker's commands at any time.

Diablo Canyon, another nuclear power plant in the US, was close to implementing a digital upgrade of its Plant Protection System by using the Triconex platform, although in the end this pilot project was canceled.



By analyzing the activity around that initiative, however, we can learn that, in fact, the US NRC was still concerned⁴⁸ about how Invensys was planning to address this known issue.

⁴⁸ <https://www.nrc.gov/docs/ML1229/ML12297A243.pdf>

7	AR (BK)	[ISG-06 Enclosure B, Item 1.16] Design Analysis Reports: The LAR does not appear to comply with the SRP (ISG-04) regarding the connectivity of the Maintenance Work Station to the PPS. The TriStation V10 platform relies on software to effect the disconnection of the TriStation's capability to modify the safety system software. Based on the information provided in the LTR, the NRC staff determined that the Tricon V10 platform does not comply with the NRC guidance provided in ISG-04, Highly Integrated Control Rooms—Communications Issues, (ADAMS Accession No. ML083310185), Staff Position 1, Point 10, hence the DCPD PPS configuration does not fully comply with this guidance.	Closed	Drafted RAI # 17 &18 to obtain an answer / report to address this topic.		(Kemper 4-12-12) Response received April 2, 29, 2012. Staff reviewed this item and still need additional information to close this item. The staff will need to review this item further	
October 16, 2012							
DCPD PPS Open Item Summary Table							
Page 11 of 76							
No	Src/RI	Issue Description	P&GE response:	Status	RAI No. (Date Sent)	RAI Response (Due Date)	Comments
		In order for the NRC staff to accept this keyswitch function as an acceptable deviation to this staff position, the staff will have to evaluate the DCPD PPS specific system communications control configuration—including the operation of the keyswitch, the software affected by the keyswitch, and any testing performed on failures of the hardware and software associated with the keyswitch. The status of the ALS platform on this matter is unclear at this time and will be resolved as the ALS LTR review is completed.					during an NRC audit at the Invensys facility. All the items noted below will be the scope of the audit.

It should be noted that Oconee uses Triconex for its turbine control system⁴⁹.

Now let's get back to TXS where we find a similar situation. The processing modules also have 4 different modes of operation⁵⁰,

1. Operation

This is the normal operating mode for cyclic processing of the logic (function diagram group -FDG- modules).

2. Parameterization

This mode is the same as the 'operation' mode, but also enables the parameterization of the logic. This means that it is possible to adjust set-points, calibration data, or other settings that are designed to be changed during regular operating conditions of the plant.

3. Test

This mode is used for functional testing. The FDG-modules can be processed in single step mode, and all external input signals can be inserted from the service unit. As opposed to the previous modes, the cyclic processing of the logic is stopped.

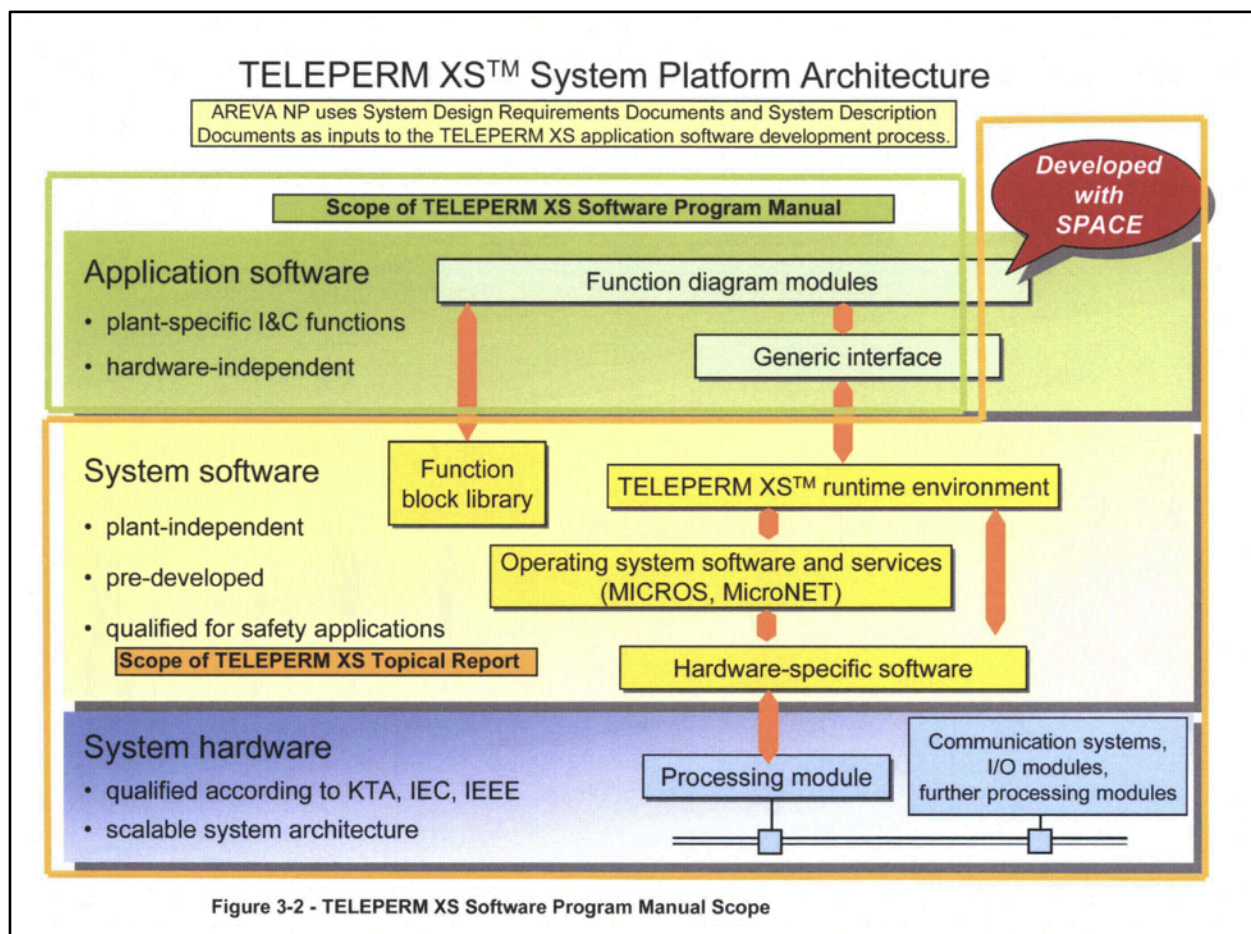
⁴⁹ <https://www.powermag.com/top-plants-oconee-nuclear-station-oconee-county-south-carolina/>

⁵⁰ <https://www.nrc.gov/docs/ML0037/ML003711856.pdf>

4. Diagnosis

In this mode, direct memory access is granted to the service unit. Special diagnosis programs can be downloaded from the service unit into RAM, which can be a valuable functionality to compromise the unit. Logic is stopped.

These modes are implemented in software, more specifically by the Runtime Environment (RTE) component, which is part of the System software. For reference purposes, the following image (AREVA) shows the system architecture of the TXS function modules (SVEx).



So, as it would happen for Diablo Canyon years later, during the review of the License Amended Request (LAR) for the new Oconee digital Plant Protection System, the NRC staff asked⁵¹ Areva/Duke how they were planning to address this issue.

⁵¹ <https://www.nrc.gov/docs/ML0823/ML082330618.pdf>

40. Staff position # 10 states: "Safety division software should be protected from alteration while the safety division is in operation. On-line changes to safety system software should be prevented by hardwired interlocks or by physical disconnection of maintenance and monitoring equipment." This section goes on to state: "A workstation (e.g. engineer or programmer station) may alter addressable constants, setpoints, parameters, and other settings associated with a safety function only by way of the dual processor/shared memory scheme described in this guidance, or when the associated channel is inoperable. Such a workstation should be physically restricted from making changes in more than one division at a time. The restriction should be by means of physical cable disconnect, or by means of keylock switch that either physically opens the data transmission circuit or interrupts the connection by means of hardwired logic."

The staff understands that the Areva TXS system proposed for use at Oconee uses key switches that sets a bit within memory, and does not use a physical disconnect. This would appear to conflict with staff guidance. Please explain why the TXS system should be approved. Include in the discussion any additional protection which may exist, or what additional protection will be provided by Oconee, beyond the normal key protection, sign-out requirements, and administrative activities which would also be used for a physical disconnect key switch.

Eventually, the NRC conducted a thorough audit⁵² of the design Areva proposed. Although that report does not seem to be publicly available, the 'S430 Module Failure' detail highlighted in the figure below, can help to deduce the implementation.

3.0 First-of-a-Kind Implementation Audit

NRC conducted an audit of the TELEPERM XS Service Unit interface design for the Oconee Reactor Protection System/ Engineered Safeguard Protection System design at

Security-Related Information Withheld - Public Version

AREVA NP Inc.

ANP-10272

Software Program Manual for TELEPERM XS™ Safety Systems
Topical Report

Revision 3

Page C-42

the AREVA NP facility in Alpharetta, Georgia on April 15 – 17, 2009 (Reference 48).

The audit was based on the presentation information submitted to NRC in Reference 47.

The failure modes and effects for the failures affecting the Parameter Change Enable Application Software logic were assessed for implications on control of undetected access to the safety processors or undetected changes of processor modes. The following conditions were evaluated:

- Failure of the Parameter Change Enable key switch
- S430 Module Failure

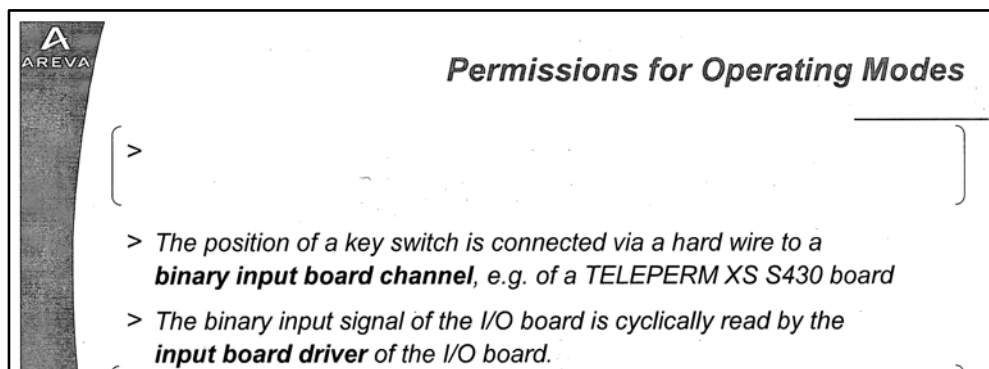
⁵² <https://www.nrc.gov/docs/ML1129/ML11294A397.pdf>

In the Triconex topical report, it was stated that the key switch ended up modifying 2 bits in memory, enough to represent each one of the 4 different operating modes of the controller. However, in the TXS case, we have seen that the documentation states that the key switch just sets one bit, but the RTE still supports 4 operating modes. So, what's going on?

The Ocone "Technical Specification Bases Change" document⁵³ gives us the answer.

BASES	
BACKGROUND (continued)	<p><u>Parameter Change Enable Mode</u></p> <p>Parameter Change Enable Mode allows each RPS instrument input channel processor to be placed in different operating modes through the use of the Parameter Change Enable keyswitches and commands from the Service Unit. Each protective channel has a keyswitch located in that channel's cabinet pair.</p> <p>Placing RPS Channels A, B, or C in Parameter Change Enable Mode through the use of the "Parameter Change Enable" keyswitch will also place the corresponding ESPS Channels A1, B1 or C1 in Parameter Change Enable Mode.</p> <p>When a keyswitch is placed from the normal Operating Mode position to the Parameter Change Enable Mode position:</p> <ul style="list-style-type: none"> • The processors continue with normal operation. • A permissive is provided that allows the Service Unit to be used to change the operating mode of the processors associated with that keyswitch.

In the TXS platform the key switch does not directly change the operating mode of the processing modules, but it releases a specific signal, which is known as a 'permissive', (essentially a protection-grade interlock) that 'allows' (elaborated below) the Service Unit to issue a command to do so. This permissive is a discrete signal, that's why just 1 bit is changing. This would be confirmed by the following presentation⁵⁴:



⁵³ <https://www.nrc.gov/docs/ML1530/ML15303A003.pdf>

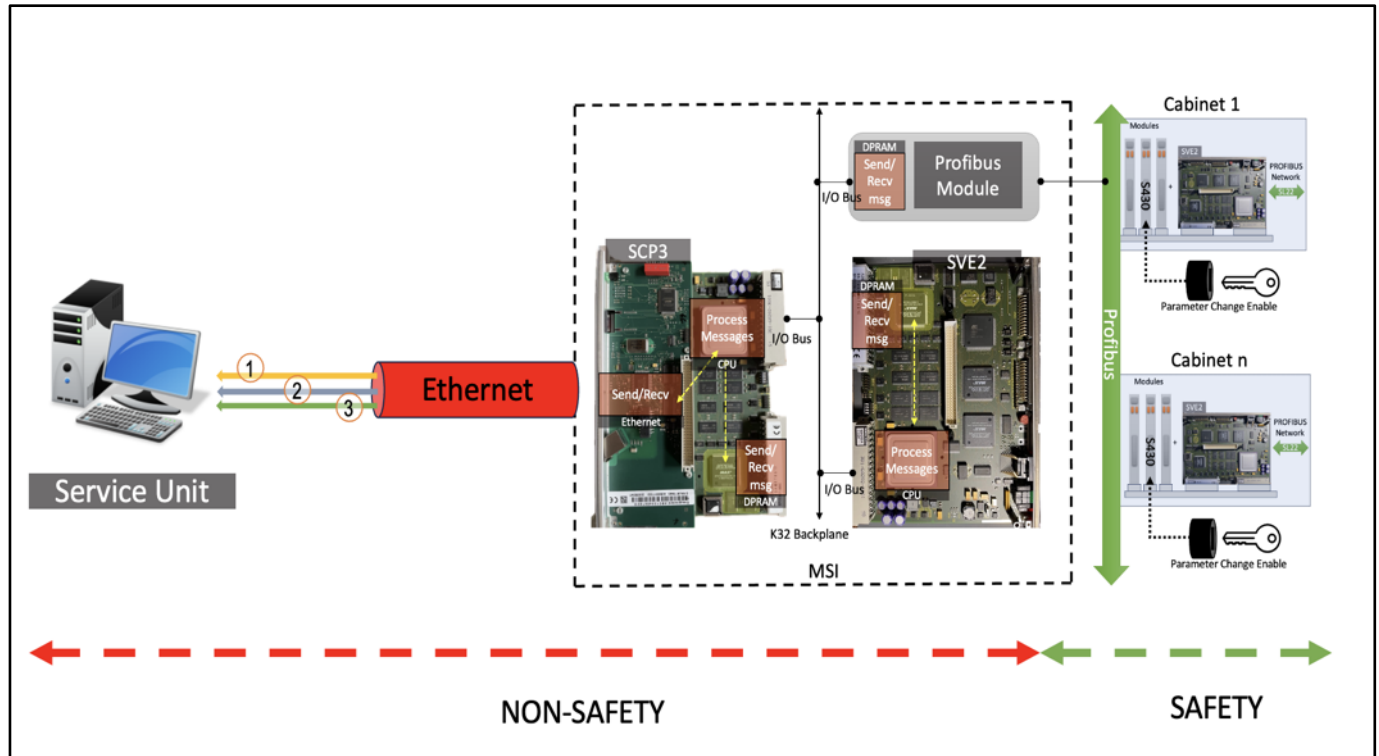
⁵⁴ <https://www.nrc.gov/docs/ML0913/ML091380436.pdf>

In Oconee there are two types of channels: instrument channels (5 RPS and 3 ESFAS) and actuation logic (8 ESFAS). These latter ones are then grouped into 2 Voters (ODD and EVEN). The Oconee functional description⁵⁵ outlines the total number of key switches, up to 10, which corresponds to this classification (5 + 3 + 2).

23.2.3 The five RPS Channel (A, B, C, D & E) PARAMETER CHANGE ENABLE Keyswitches, three ESFAS Channel (A, B & C) PARAMETER CHANGE ENABLE Keyswitches and two Voters (Voter 1 ODD, Voter 2 ODD, and the Status Computer in 1PPSCA0017 and Voter 1 EVEN, Voter 2 EVEN, and the Status Computer in 1PPSCA0018) PARAMETER CHANGE ENABLE Keyswitches for any solitary Oconee Unit use the same key code. The keys are different for each Oconee Unit. Keys shall be non-removable in the ENABLE position.

It seems there is no 1:1 correspondence between key switches and modules, but it would likely be a cabinet-wide permissive signal. If so, it would enable changing the operating modes for a specific set of processing modules.

Now, let's put everything together.



⁵⁵ <https://www.nrc.gov/docs/ML0910/ML091050333.pdf>

The communication between the different processing modules is based on three different types of messages:

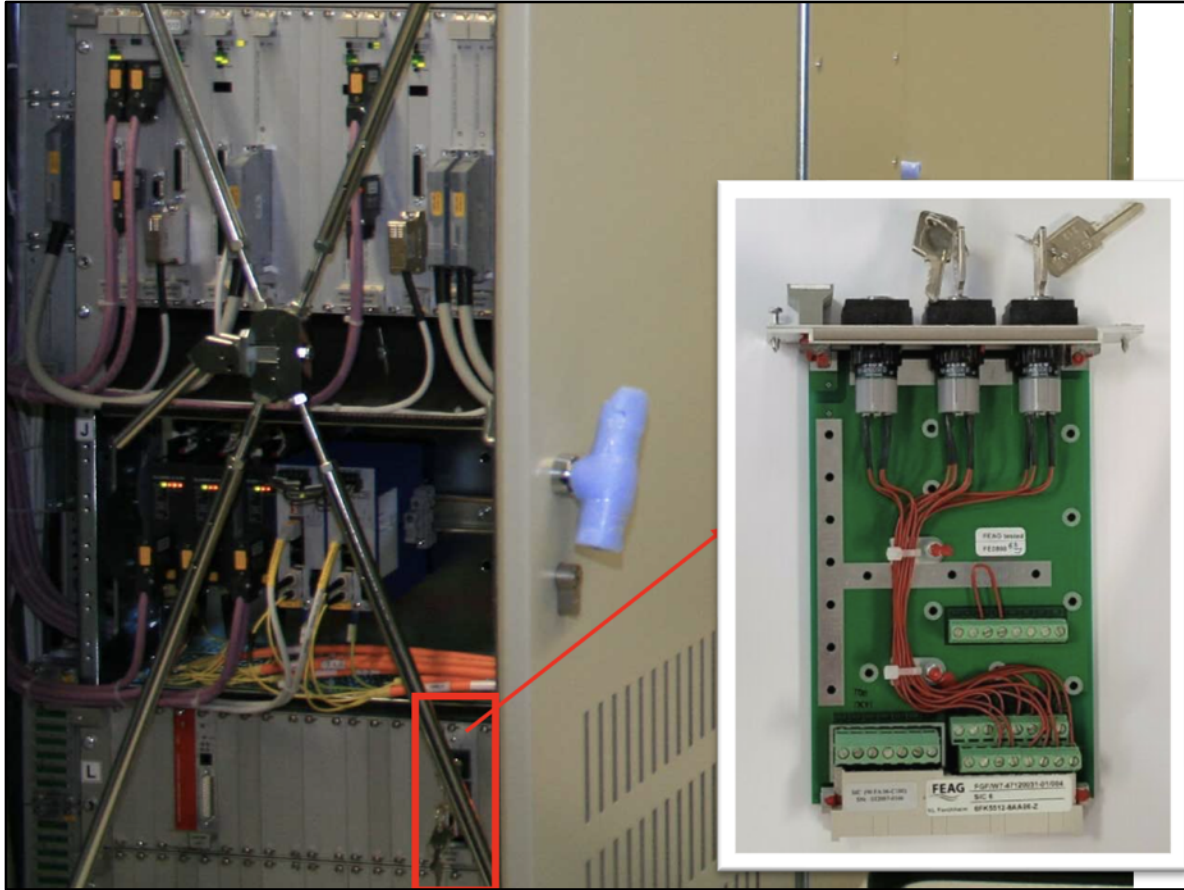
1. Control
Commands issued from the Service Unit to the various processing modules. These messages will be analyzed and validated by the MSI before being delivered.
2. Data
Signals and other related data coming in to or going out from the application software (FDG modules)
3. Signaling
These are intended to be consumed by the MSI and the Service Unit. This includes error messages, command responses, status messages, and so on.

There are three main communication paths:

- a) Ethernet
This is used between the Service Unit and the MSI, as part of the non-safety/safety communication mechanism.
- b) K32 Backplane
Dual-Port RAM (DPRAM) is used to keep the independence between the communication logic and functional logic (compliance with “Staff Position 4”⁵⁶). It is used for inter-module communications.
- c) Profibus
Profibus is used in the safety part of the network.

We can presume that each cabinet will contain a TXS key switch module, such as the SC61, to implement the ‘PARAMETER CHANGE ENABLE’ permissive.

⁵⁶ “The communication and function processors should operate asynchronously, sharing information only by means of dual-ported memory or some other shared memory resource that is dedicated exclusively to this exchange of information.” <https://www.nrc.gov/docs/ML0833/ML083310185.pdf>



The status of this signal is both transmitted to the Operator Aid Computer (OAC) via the TXS Gateway, and also apparently hardwired to the alarms panel (statalarm) in the MCR. Additionally, each processing module cyclically transmits its current operating mode to the MSI, information that is consumed by the Service Unit.

Therefore, the steps to, legitimately, change the operating mode of a processing module would be as follows:

- Physically use the key switch in the corresponding cabinet to release the “PARAMETER CHANGE ENABLE” permissive.
- Use the Service Unit to issue a command that changes the operating mode for the specific processing module. This means that once the “PARAMETER CHANGE” permissive is released, changing the operating mode to ‘TEST’ or ‘DIAGNOSIS’ wouldn’t require any further physical interaction so it can be done entirely via software from the Service Unit.

The PARAMETERIZATION Mode allows changes to specific parameters or performance of tests from the GSM screens. Permission to change from the OPERATION mode into the PARAMETERIZATION mode is provided by the Parameter Change Enable Keyswitch. After the permissive is provided from a

system processor via its Keyswitch, communication from the Service Unit to that processor is allowed to change its operating mode. Placing the PROCESSOR into the FUNCTION TEST and DIAGNOSTIC modes requires first enabling the PARAMETERIZATION mode with the keyswitch and then setting a separate parameter to enable these modes with the GSM.⁵⁷

Please note that although we have studied the implementation for Ocone, due to the amount of information available, the same general architecture applies for the common design of other TXS-based Plant Protection Systems. However, there may be some differences among implementations, such as:

- A requirement to implement a different key switch that must be positioned to the specific operating mode, instead of just releasing a permissive and using the SU to change the operating mode.
- A requirement to use a centralized key switch in the MCR to release the 'PARAM' permissive.

In any case, what is important is the fact that a mechanical element is required to change operating modes. Therefore, a malware seeking to infect the TXS processing modules would need to wait until detecting this permissive(s), which represents a mitigation. However, once a TXS module has been compromised, this security barrier fades away as the operating-mode permissives are discrete signals exclusively handled by software. A key switch that does not implement either a physical disconnection or a hardware interlock, lacks any effectiveness.

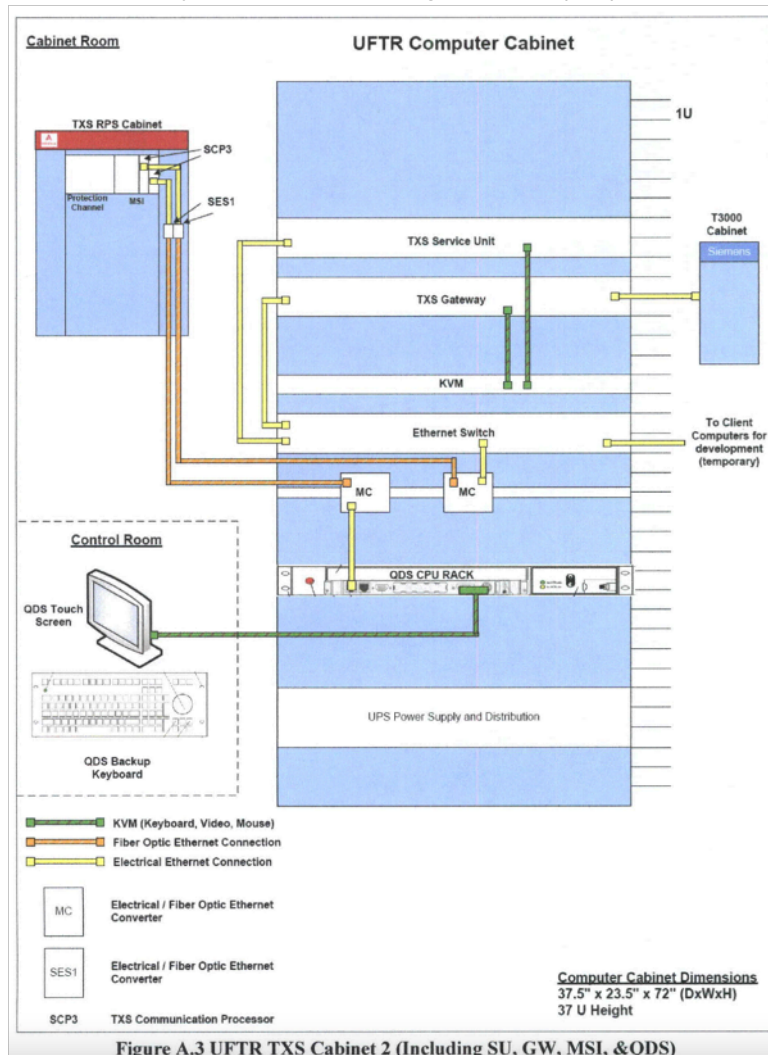
So, in view of this scenario, and its similarity with Trisis, one of the obvious questions one may raise would be if a Trisis-like operation would have any chance to work against TXS too. I'm sure about the answer so I'll explore this scenario in the next main section.

⁵⁷ <https://www.nrc.gov/docs/ML2018/ML20189A086.pdf>

3.3.3 - Research Reactors

TXS is not only used in commercial NPPs but also in research reactors, such as:

a) The University of Florida Training Reactor (US)⁵⁸



In this architecture we can clearly identify some of the TXS components that I found available on eBay:

- The SES1 device
- The Teleperm XS Qualified Display System, consisting of the QDS-CPU (located in the QDS-SubRack) and the QDS Touchscreen.

⁵⁸ <https://www.nrc.gov/docs/ML1019/ML101950366.pdf>

These QDS components comply with the safety requirements to be used as part of a 'Post-Accident Monitoring System' (PAMS). Depending on the architecture, we can find either unidirectional or bi-directional communication between the QDS-CPU and the MSI, so it is also an attack vector to consider. Any kind of temporary physical access to the QDS Sub-RACK allows one to trivially compromise the QDS-CPU in multiple ways.



Figure - QDS-CPU - front panel (USB,PS/2 Mouse/Keyboard, RS232, VGA, and RJ45 ports)

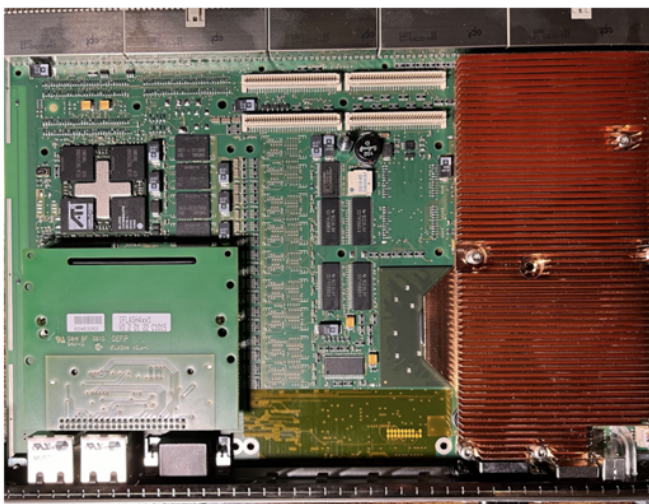


Figure - QDS-CPU main board (iFlash IDE board at bottom left)



Figure - iFlash Board

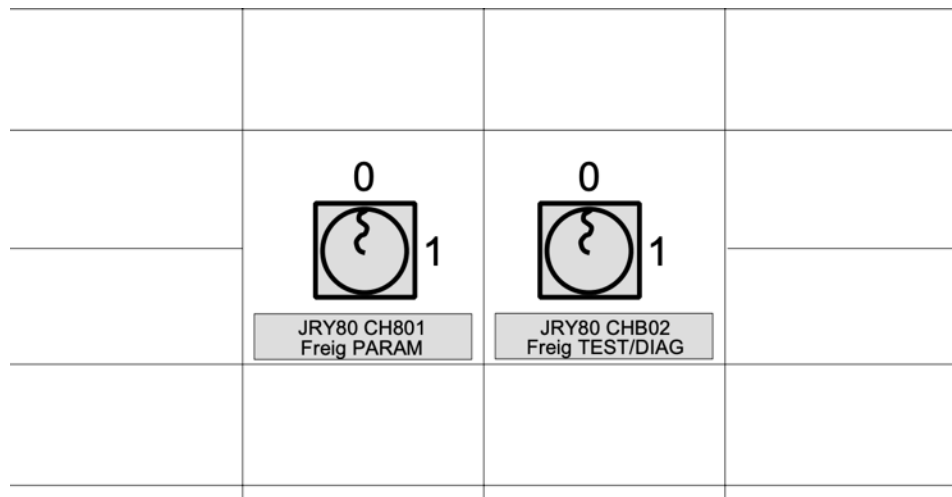
b) AKR-2 - Technical University of Dresden (Germany)

The AKR-2 is not only a real student attraction at the TU Dresden, but is also attractive for the qualification and further training of employees from the nuclear industry or other specialist institutions in the field of reactor physics, reactor control technology and radiation and neutron flux measurement technology, not least because it is equipped with the TELEPERM XS digital control technology system.⁵⁹

⁵⁹ https://tu-dresden.de/ing/maschinenwesen/ifvu/wket/ausbildungskernreaktor-akr-2?set_language=en



There is something interesting to note in the Control Desk of the AKR-2⁶⁰. As we anticipated, there may be different variations to implement the change of operating modes in the TXS processing modules. In this case, in the conventional operation and signal panel (the panel on the left side of the control desk) we can find two different key switches, one for the 'PARAM' mode and another different one for enabling 'TEST/DIAG' modes, which provides another layer of protection against malware-based attacks (and unintended maintenance operations).



⁶⁰ https://tu-dresden.de/ing/maschinenwesen/ifvu/wket/ressourcen/dateien/akr2/Lehrmaterialien/aufbau_e.pdf?lang=en

3.4 Trisis-like attacks against the TXS platform

One of the inherent (non plant-specific) characteristics of the Teleperm XS (TXS) platform is the lack of a cryptographically secure method to verify any new firmware. The TXS system software (loaded in SCPx, SVEx modules) uses CRC32 to validate the integrity of the loaded firmware/configuration.

Obviously, this is an issue as there is no way a CRC can be used to validate the authenticity of the firmware. For an n -bit CRC whose polynomial is of length $n+1$, we just need to modify, at most, n bits in the target file to transform the current CRC into the desired CRC. This means that a malicious actor can backdoor legitimate TXS firmware files (cpuxxxx.mic files) in such a way that their computed CRC will match the expected value. This is feasible even when several CRCs are calculated from the same file due to the linearity of CRC calculations.

The Service Unit can be used to request those CRC values from the processing modules to compare them with the local versions stored in the Service Unit.

Лист № 7
Всего листов 29

Таблица 1 - Идентификационные данные ПО ИС-УСБ-АЭС-TXS-440

Наименование ПО	Идентификационное наименование ПО	Номер версии (выпуска) ПО	Цифровой идентификатор ПО	Алгоритм вычисления цифрового идентификатора
Пакет программного обеспечения TXS	TXS CORE Software	3.3.x	Номер версии (выпуска) ПО	Номер версии (выпуска) ПО считывается с экрана терминала сервисной станции
Файлы с прикладными программными компонентами модулей обработки SVE2 шкафов ПТК АЗ-АЗ УСБТ конкретного энергоблока АЭС	cpu1311.mic	Не используется	Приведен в эксплуатационной документации на ПТК АЗ-АЗ УСБТ конкретного энергоблока АЭС (см. примечание)	CRC32
	cpu1415.mic			
	cpu1610.mic			
	cpu2311.mic			
	cpu2415.mic			
	cpu2610.mic			
	cpu3311.mic			
	cpu3415.mic			
	cpu3610.mic			
Примечание – изменения цифровых идентификаторов ПО модулей обработки SVE2 в процессе эксплуатации проводят при условии положительных результатов отладки, верификации и тестирования всех внесенных в ПО изменений и отражают в эксплуатационной документации ПТК АЗ-АЗ УСБТ				

Figure - Identification of TXS firmware files for Russia's Kola NPP - VVER-440 reactor

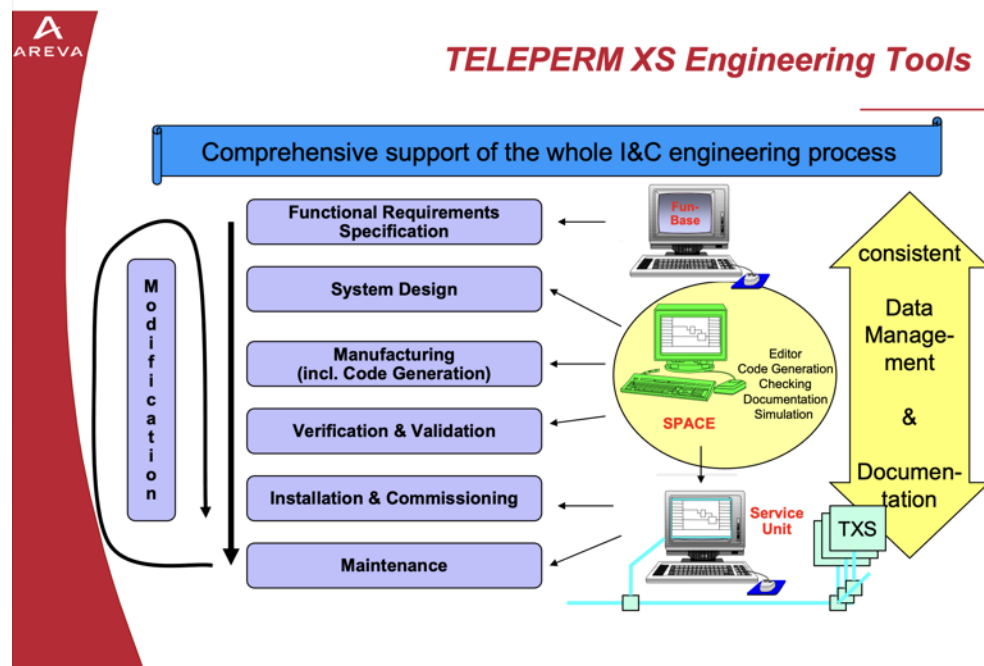
As a result, backdooring a firmware file will always be an option, although probably not the best one, as it would require a very precise approach to inject the malicious firmware at the right moment in the supply chain. There are several reasons that make this approach especially difficult in this context.

- There are different verification methods intended to ensure the integrity, from the functional perspective, of the firmware. These procedures are based on the static analysis of the compiled firmware by using reverse engineering.

For instance, MALPAS⁶¹, a static analysis toolset, is used in the UK to verify the TXS firmware that will be deployed at Hinkley Point C NPP⁶².

This approach is also aimed to catch compiler-induced bugs, which is something to consider during the development of safety software, not only from the functional perspective but also for security purposes. For instance, I found an interesting example (PCIE.dldd: RESET_MIB_DATA IOCTL Double fetch⁶³) of this type of issue while reverse engineering a safety-critical avionics firmware. In a kernel-mode driver, the developers failed to declare as 'volatile' an attacker-controlled (user-mode) variable that was then used in a switch statement. While in the source code this may look correct, internally the C compiler generated the code to use a double fetch of the offending variable instead of keeping it in a register, which caused an exploitable race condition.

- Firmware is extensively tested before being deployed (Verification & Validation). Additionally, there are well defined procedures to deploy firmware to the TXS processing modules, certain validations are performed immediately prior to installing the files.



⁶¹ https://en.wikipedia.org/wiki/MALPAS_Software_Static_Analysis_Toolset




⁶² <https://www.atkinsrealis.com/en/media/trade-releases/2020/2020-04-23>

⁶³ <https://act-on.ioactive.com/acton/attachment/34793/f-5fe563fb-fc4c-4ce5-93bd-e20a154b300b/1/-/-/-/IOA-Avionics-Research-22-Santamarta.pdf>

Despite these constraints, backdooring a firmware would still be a viable approach for attacking TXS-based deployments if the actors are in a position to compromise specific assets, (e.g. TXS maintenance laptop).

However, as previously introduced, the most promising target is the Service Unit. The attack surface is also relevant, as it includes both COTS and open-source software.

TELEPERM XS SYSADMIN-Linux Fundamentals

DURATION	LOCATION	LANGUAGES
2 days	 Framatome Karlstein	  German / English
<div> <div> TARGET GROUP This course is intended for technicians responsible for the administration of a TELEPERM XS system </div> <div> OBJECTIVES Upon successful completion of this course, participants will be able to: <ul style="list-style-type: none"> - State the basic functions of the TELEPERM XS Service Unit (SU) - Configure TELEPERM XS hardware and software for a SU - Install and test TELEPERM XS Core Software - Perform basic administrative tasks on a TELEPERM XS Service Unit </div> </div> <div> CONTENT The course is based on the Linux operating system and TELEPERM XS software version higher than 3.3. using the TELEPERM XS Service Unit. The participants learn how to set up and administrate users, groups and printers. Furthermore, they learn how to install TELEPERM XS software packages. YaST system administration is also dealt with. Handling of the KDE desktop environment will be consolidated. The participants consolidate the acquired knowledge in practical exercises. The course covers the following topics in detail: <ul style="list-style-type: none"> - Overview of TELEPERM XS and Linux - Installation of SUSE Linux Enterprise Server - KDE and Linux concepts - Creation of TELEPERM XS users and groups - Installation of TELEPERM XS software packages - Configuration and administration of TELEPERM XS - YaST Control Center - Practical exercises, including testing of the installation </div>		

64

I'm omitting from the scope of this research the methods that could be used to gain initial access to the Service Unit. I'm assuming that those actors in the position to carry out a cyber-physical attack against a nuclear reactor, will have the resources, both human and technical, to achieve this crucial objective. However, it's worth discussing 'when' this infection may happen.

⁶⁴ https://www.framatome.com/solutions-portfolio/docs/default-source/default-document-library/bu-documents/ic-fram_23_002_worldwide_training_catalog_interactive.pdf

3.4.1 - The 'When'

Obviously, there is an inherent risk in having a computer with the ability to control the RPS & ESFAS, permanently connected to the network. However, if disconnected, a significant amount of information for maintaining the situational awareness of the plant may then be missing.

This has created an interesting situation that apparently raised some concerns within the US NRC staff⁶⁵ during the review of the U.S. EPR design.

Question #3

➤ NRC Question:

- ◆ Some European regulators are saying the service unit is not permanently connected in their designs. AREVA says it is permanently connected for all EPRs at last public meeting, but would verify. If the service unit is temporarily connected in the European EPR designs, how is it technically being implemented?

➤ AREVA NP Response:

- ◆ At the July 21 meeting, the question was verbally posed to AREVA with reference to France (i.e., FA3).
- ◆ FA3: service unit is permanently connected.
- ◆ OL3: service unit is permanently connected.
- ◆ In the U.K., it has been agreed with HSE (September 2009), that the service unit will be disconnected during plant operation except when needed. It will simply be disconnected or powered down and will be connected when needed to maintain the system. Technical details of implementation have not been finalized.

Basically, there is no clear picture of which NPPs keep the SU permanently connected and which do not. For instance, according to Areva, Flamanville 3 (France), and Olkiluoto 3 (Finland) would be operating with the SU permanently connected.

Although possibly not permanently connected, it is guaranteed that the Service Unit will be connected periodically, due to the requirements of the plant: maintenance tasks, adjust calibration data and setpoints, verification of parameters, and so on. As is done, for instance, at Oconee⁶⁶:

The setpoints in the software are manually verified every 92 days. The protective channel interposing relays are manually actuated every 92 days. RPS logic is reverified every

⁶⁵ <https://www.nrc.gov/docs/ML1024/ML102460343.pdf>

⁶⁶ <https://www.nrc.gov/docs/ML2018/ML20189A086.pdf>

refueling outage by rebooting the channel computer and checksums are verified at that time.

From the previous paragraph we can surmise a prominent period when the opportunities to compromise the Service Unit, and subsequently certain TXS processing modules, significantly increase: refueling outages.

A refueling outage is a big event for nuclear power plants, bringing hundreds of additional workers from different fields (including contractors), usually taking place over weeks. For example:

To support the refueling outage, approximately 2,000 additional workers will travel to Calvert for several weeks, filling nearby hotels to capacity and increasing foot traffic in restaurants and shops.⁶⁷

In addition to substituting the spent fuel, these outages are also leveraged to perform maintenance tasks, including those associated with the digital I&C systems. This crowded, busy environment is the perfect scenario for carrying out the infection. For instance, the TXS cabinets generate an alarm when opened, while this alarm may be suspicious during regular operating conditions, it might be simply overlooked during a refueling outage.

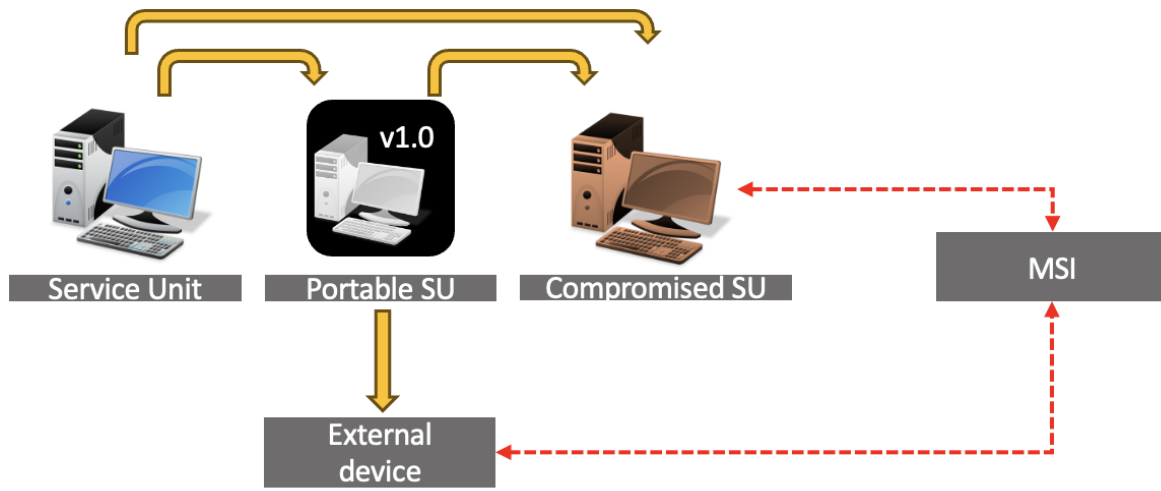
3.4.2 - The 'Why'

The rationale to consider the Service Unit as the main target could be explained from two different perspectives:

- a) Compromising the Service Unit itself.
 - This would guarantee bi-directional access to the MSI, a basic requirement for any malware-based operation. Additionally, it would allow full control of the information displayed to the technical staff.
- b) Compromising another device in the same network segment which allows forging the Service Unit's network traffic.
 - As the MSI uses the MAC Address of the Service Unit as an access control, this is a potential alternative to the first approach.

In both cases, the prospective elements involved are illustrated in the diagram below:

⁶⁷ <https://www.constellationenergy.com/newsroom/2017/calvert-cliffs-refueling-outage-powers-local-economy.html>



1. Service Unit

The legitimate Service Unit (SU).

2. Portable SU

I'm using this term to refer to a library that would enable an implant to replicate core functionalities from the Service Unit without having to rely on the Service Unit itself. For instance, changing the value of a specific setpoint and bypassing the sanity checks the SU implements. This would plausibly require significant reverse engineering efforts to understand the underlying network protocol and messages.

There is a trend where the companies manufacturing industrial controllers are specifically making significant efforts to prevent 3rd party software from communicating with their equipment. The idea is that the only legitimate communication channel should be established via the original manufacturer's engineering tool, which can act as a mitigation against implementing certain cyberattacks.

As an example, several years ago Siemens designed a new protocol, S7CommPlus, for interfacing with certain S7 controllers. The implementation of this complex cryptographic protocol is heavily obfuscated in a DLL used by the Siemens' TIA Portal engineering tool. I spent several weeks reverse engineering this protocol, eventually being able to build a small portable C library. This was possible because S7CommPlus uses some fields derived from a custom elliptic curve cryptography implementation but without binding them with other cryptographic materials also present in the protocol's handshake, thus enabling the ability to statically reuse the elliptic curve calculations from a previous connection.

It is then assumed that, even if the Service Unit implements a complex, obfuscated logic, determined actors will eventually be able to build a 'Portable SU' library that can mimic specific functionalities of the original 'Service Unit'.

3. External Device

This represents any equipment with a privileged position in the non-safety network segment (access to the MSI) with two characteristics:

- Susceptible to being compromised.
- Able to generate traffic with a spoofed MAC address.

This would potentially enable impersonating the Service Unit, as its MAC address is used by the MSI as an access control.

By having the 'Portable SU' running in this compromised device, attackers would be able to deploy cyber-physical payloads without relying on the original Service Unit.

Please note that this 'External Device' may or may not exist in an NPP. Let's remember that, ideally, non-safety equipment should communicate with a safety one only through a point-to-point link.

4. Compromised SU

This represents the original Service Unit, once it has been compromised.

The implant in the SU may be relying on the original Service Unit components (e.g. via hooking) as part of its logic, or using the 'Portable SU'. This latter option would certainly be less prone to detection by endpoint security solutions.

For instance, Trisis was found in a compromised engineering workstation using its own TS1131 library to communicate with the safety controllers.

3.4.3 - The 'How'

Once the malicious actors have positioned themselves as an 'alternative' Service Unit, the implant can just stay dormant, waiting for the right moment to either infect or attack. A simple outline of the scenario would be as follows.

The implant detects the permissive signal that marks the start of the time window where it will be possible to change the operating mode of the TXS modules.

Depending on the characteristics of the planned cyber-physical attack we have two main options:

a) Attack

The implant will wait (or force a change in the operating mode if possible) until the target TXS processor is in the right operating mode.

For instance, in a set-point based attack, when the 'PARAM' mode has been enabled, the implant will proceed to modify those set-points required to 'deploy' the cyber-physical payload. This must be carefully implemented to match the operational procedures (e.g. manual review of set-points after configuration). Depending on the capabilities and position of the implant, this may require hooking into the SU's Graphical Service Monitor (GSM, the component in charge of the GUI) component in order to control the information received by the operator.

b) Infect-then-attack

The malware will presumably gain code execution (potentially achieving persistence) in the targeted modules by first changing the operating mode to 'DIAGNOSIS' and then abusing some of the functionalities enabled in that mode. For instance, by deploying an implant (malicious application software) via the 'WRITE_FDG' message.

However, the only two operating modes that are able to run logic in the TXS processing modules are 'OPERATION' and 'PARAM'. So, in terms of RAM-only persistence, as in Trisis, it should be noted that changing back to 'OPERATION' or 'PARAM' after setting 'TEST' or 'DIAGNOSIS' modes theoretically requires a processor reset. Ideally, this is something the module's implant should explore to keep the infection process as stealth as possible.

4. Cyber-Physical Attacks

It is within the realm of possibility... Unlikely..., but possible... If all the remote-controlled devices fail, then, faced with the threat of the pile exploding, someone would have to climb through this hatch and manually insert the rods of the cadmium moderators into the graphite.

Stanislaw Lem. *The Astronauts* (1951)

I would like to introduce this section by clearly defining its scope and limitations to avoid any kind of misunderstanding.

What is being studied herein is a methodology intended to approach the, fictitious, mandate of forcing a nuclear reactor into a mode of operation that goes beyond its safety limits, in order to trigger a specific sequence of physical events, exclusively via ‘cyber’ means, that might result in fuel damage, core melt and/or release of radioactive materials.

— Disclaimer —

It must be clear that this is a theoretical exercise, although based on realistic assumptions, technical facts, official documentation, commercial software used for severe accident simulations and reverse engineering of digital safety I&C hardware components.

I am not claiming that any of the scenarios elaborated herein may actually work in the real world, under any circumstance. I am not claiming either that this methodology represents a real threat for any of the referenced nuclear power plants or currently commissioned PWRs.

Finally, I highly discourage the use of any partial, distorted or biased interpretation of this research to create sensationalism against nuclear energy based on unscientific reasoning and FUD.

4.1 Introduction

Cyber-attacks against the safety systems of nuclear reactors can be, ideally, designed to inflict different types of damage, depending on the requirements, objectives and limitations of the operation. I use the term “ideally”, because due to the complexity of the physical processes involved, it is extremely difficult to ensure that the resulting damage, if any, will eventually match what has been initially planned, by excess or defect.

Let’s define four types of scenarios, according to both the perception that such an attack would have on the population as well as its measurable damage.

A) Disaster

The International Nuclear and Radiological Event Scale⁶⁸ can be used to describe what is required for a successful cyber-attack to provoke a 'disaster'. Any death or release (even minor) of radioactive materials into the environment would likely be perceived as a disaster.

In order to achieve this outcome, the attackers need to provoke a failure of digital, analog, human and physical barriers in a short period of time, which is extremely unlikely to occur in any modern PWR NPP.

The truth is that from an offensive perspective, it is orders of magnitude easier to attack other kinds of industrial facilities (especially chemical plants), to cause a much worse outcome. However, due to the public perception of nuclear energy, there is something highly symbolic in targeting an NPP. So, the implications of a successful attack of this type would surely be devastating, in a wide sense.

B) Destruction

This kind of attack would be aimed at destroying specific components or structures located inside the Reactor Pressure Vessel (RPV). This situation would certainly guarantee a long-term (even permanent) outage of the targeted unit(s) and most probably the entire NPP.

A successful attack of this kind does not necessarily require provoking a failure of all the barriers previously laid out. There are certain specific actions that may plausibly compromise the structural integrity of certain important components, not even essential ones. However, we must bear in mind the massive post-incident efforts that would need to be undertaken to ensure the safety and security of the targeted NPP.

C) Disruption

A botched attempt to attack the safety systems of a PWR would likely trigger either an automatic or an operator-initiated reactor shutdown. On the other hand, the cyber-physical attack may actually have a controlled reactor shutdown as its primary goal. For instance, the UK includes this scenario in its National Risk Register⁶⁹.

⁶⁸ <https://www.iaea.org/resources/databases/international-nuclear-and-radiological-event-scale>

⁶⁹ https://assets.publishing.service.gov.uk/media/64ca1dfe19f5622669f3c1b1/2023_NATIONAL_RISK_REGISTER_NRR.pdf

Cyber attack: civil nuclear

Civil nuclear power is of strategic importance to the UK's energy security and net zero ambitions and in turn, must continue to strengthen its resilience to dynamic and evolving cyber threats. Cyber security in the civil nuclear sector is managed through a combination of nuclear safety and security regulatory requirements, a defence-in-depth approach and sector-wide collaboration under the 2022 Civil Nuclear Cyber Security Strategy. The combination of these approaches drives a holistic and robust risk mitigation on cyber.

Scenario

This scenario assumes a cyber attack that could require a controlled shutdown of a civil nuclear generating site as a protective measure. This could result in a temporary loss of supply to the UK National Grid until its restoration or generating capacity could be increased elsewhere. Impacts from this loss could vary depending on how power redistribution is managed.

Response capability requirements

The National Grid requires the capability to restore grid systems and manage power distribution. Local Resilience Forums are required to manage potential regional-level impact to essential services as part of their arrangements for managing disruptions from loss of power. Functional back-up generators would be required for a range of other critical infrastructure sectors to reduce impact on essential services.

Recovery

The reactor's return to service could be a lengthy process, depending on the nature of the incident, while replacements and repairs take place due to strict regulatory controls designed to ensure nuclear safety and security.

This risk is featured in the full matrix on page 15, representing the averages of multiple different scenarios presented together in the 'cyber attacks on infrastructure' category.

However, it should be noted that a controlled reactor shutdown is an absolutely normal protective action, so the attackers could trip the reactor much more 'easily' by going after specific non-safety systems. For instance, the turbine control system which is located in the conventional part of the NPP.

In any case, either a failed or a "first warning"-type cyber-physical attack would likely lead to a medium-term disconnection of the entire NPP, until a thorough forensic investigation can determine that it is safe and secure to initiate operations again.

D) Deception

I consider 'deception' as an essential part, if not the primary objective, of a cyber-physical attack against nuclear reactors, or nuclear facilities in general terms.


Radioactivity is invisible for the human eye. Therefore, we necessarily rely on instruments, devices and all kinds of equipment to assess the impact of an incident involving potential releases of radioactive materials. The problem is that this information can be spoofed⁷⁰, thus drawing a picture of the scenario of an attack that does not correspond to the actual physical conditions.

⁷⁰ <https://www.iaea.org/projects/crp/j02017>

I demonstrated how to perform ⁷¹ this kind of attack in BlackHat USA 2017. Then during the beginning of the Russian invasion of Ukraine, the radiation spikes allegedly detected in the Chernobyl Exclusion Zone became a vivid example of this scenario⁷². To say it in soft terms, almost nothing of what was published in the media about the potential risks for Chernobyl during that time was physically sound. Unfortunately, as usual in any nuclear related news coverage, FUD prevailed over facts.

In 2023, after analyzing the radiation measurements, and reversing radiation monitoring devices and related software I presented research⁷³ demonstrating the plausible fabrication of those spikes. Shortly after, an international mission of nuclear experts, led by Ukraine with funds from the Norwegian nuclear authority, confirmed⁷⁴ the “abnormal origin” of those radioactivity readings. They also explicitly amended the initial official explanation, which pointed to the resuspension of radioactive dust due to traffic of Russian military vehicles, as ‘barely plausible’⁷⁵.

Preliminary assessment of the radiological consequences of the hostile military occupation of the Chornobyl Exclusion Zone

Yu Balashevskaya^{6,1} , M Chala¹ , Z Ivanov¹ , A Myshkovska², I Shevchenko¹ ,

O Pecherytsia¹ , Y Yesipenko¹ , K Siegen³, L Jova Sed⁴, G Smith⁵  [Show full author list](#)

Published 25 September 2023 • © 2023 The Author(s). Published on behalf of the Society for Radiological Protection by IOP Publishing Ltd

[Journal of Radiological Protection](#), Volume 43, Number 3

On 24 February 2022, readings up to 500 µSv h⁻¹ compared with the usual values of 0.10 µSv h⁻¹ which cannot be explained by resuspension. The absence of additional radiation contamination in the vicinity of some ARMS posts (Opachychi, Straholissya and Kupuvate) and in some other locations in the ChEZ indicates the abnormal origin of the DER measurement results by dosimeters of the ARMS of the ChEZ from 24 to 25 February 2022. The theory that abnormal DER measurements were caused by significant dust formation due to disturbance of the topsoil by heavy military equipment is barely plausible.

As a result of this we can see that, you can actually attack a nuclear reactor, or just pretend that you have attacked a nuclear reactor. In both cases it will take a significant amount of time to clarify things, if they ever become clear at all. In the meantime, the actors behind the operation may leverage the situation for their own gain.

⁷¹ <https://www.blackhat.com/docs/us-17/wednesday/us-17-Santamarta-Go-Nuclear-Breaking%20Radition-Monitoring-Devices-wp.pdf>

⁷² <https://www.wired.com/story/chernobyl-radiation-spike-mystery/>

⁷³ <https://www.reversemode.com/2024/01/what-really-happened-in-chernobyl.html>

⁷⁴ <https://www.reversemode.com/2024/06/ukraines-nuclear-regulator-confirms.html>

⁷⁵ <https://iopscience.iop.org/article/10.1088/1361-6498/acf8d0>

The 'Deception' component can also act as a complement for the other D's (Disaster, Destruction and Disruption) to either make up for a failed attempt or to maximize the impact of a successful attack.

4.2 Preparation

Carrying out a cyber-physical attack against a PWR requires a significant amount of time and resources just to decide what would be the best way to do so.

There are several factors that influence this decision-making process, to name a few:

1) Level of damage

There would be significant differences in the way the operation is approached when the mandate requires provoking a Disaster (as in the previous section) versus just a Disruption.

2) PWR design

Although all of them share the same physical and operational principles, not every PWR is designed in the same way. Therefore, the design of the target PWR needs to be carefully studied, because there are multiple elements that may directly impact how the cyber-physical attack unfolds: elevation of the components (SGs, pressurizer, etc.), core geometry, capacity of safety and relief valves, structures, systems and components (SSCs).

3) NPP design

In the same line of the previous factor, the NPP design as a whole needs to be taken into consideration. Operational procedures, defense in depth and diversity implementation, mitigating systems and strategies, design of the control room, and so on.

4) Country

The country where the target is located also plays an important role. It is important to consider its regulations, how the operators are trained, its geographical location (e.g. is it close to any allied country of the actors carrying out the operation? If so, a large radioactive release can hypothetically impact them as well.), as well as its cultural characteristics.

5) Load Following vs Baseload

Certain NPPs are licensed to operate in load-following mode instead of baseload. This implies that there are important differences in the way that an NPP will be operated, including its related digital I&C systems (especially the Control Rods Drive Mechanism).

6) Mode of operation

An NPP can be found operating in several modes: Power Operation, Startup, Hot standby, Hot shutdown, Cool shutdown, and Refueling. Each of these modes present specific opportunities to carry out an attack. However, Power Operation is naturally the primary target.

7) Fuel cycle

Attacking at the Beginning of Cycle (BOC), at the End of Cycle (EOC) or in between will depend on the characteristics of cyber-physical attack. For instance, if it's a reactivity-focused attack (e.g. boron dilution) then BOC will be preferred, as the fresh fuel will provide excess reactivity. On the other hand, at the EOC, as the fuel is now depleted, we will have a significant amount of fission products which can certainly increase the impact of those attacks aimed at releasing radioactivity.

Here follows what I would consider the minimum set of documentation to be consumed during the preparation stage for a cyber-physical attack:

- Safety Analysis Report (SAR)
- Safety Evaluation Report (SER)
- Technical Specifications
- Limiting Conditions for Operation (LCO)
- Trip reports
- Topical Reports
- Licensee Event Reports (LERs)
- Training materials and operator manuals

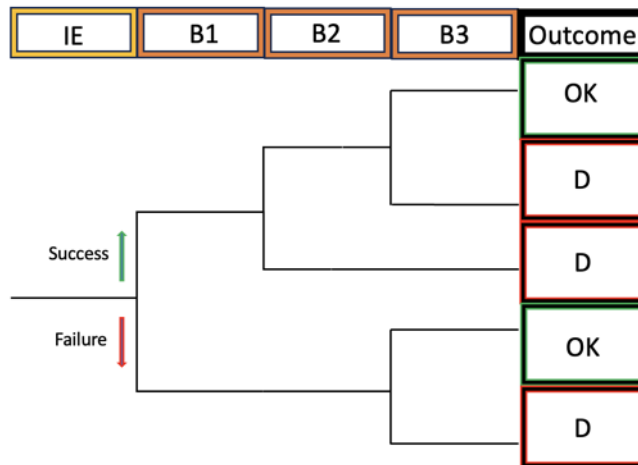
4.3 Implementation

From the perspective of the attackers, there is good and bad news. The good news is that all the Probabilistic Risk Assessment⁷⁶ (PRA) work behind the licensing of a NPP can be repurposed to find the weakest points in the target. The bad news (good news for everyone else) is that commonly these safety analyses are so thorough that there are no obvious defects to be found.

Essentially the cyber-physical attack can be characterized by a mix of inductive (events tree) and deductive (faults tree) reasoning. These two analytical approaches are widely used in the nuclear sector, so we can leverage the work done to come up with an optimal adversarial version of them.

Essentially, the idea is to find the right sequence of physical events that can escalate into a severe accident, bypassing the different Defense in Depth and Diversity (D³) systems, and thus eventually causing the desired damage. We can represent this in a sample event tree.

⁷⁶ <https://www.nrc.gov/about-nrc/regulatory/risk-informed/pra.html>

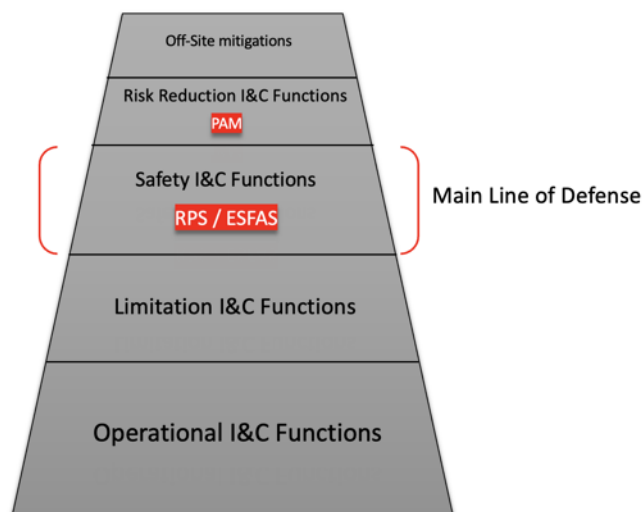


Where

- IE is the initiating event, which will trigger the sequence of physical events required for the attack.
- B1, B2, B3...Bn are the different Barriers (that can be structures, systems and/or components) that conform to the D³ implementation of the target NPP.
- Outcome represents the final state: 'OK' for an unsuccessful attack and 'Damage' (D) when it succeeded.

It is important to note that passive safety systems are out of the scope of this event tree, but they represent an important, ineludible barrier that ultimately mitigates any attempt to push the reactor out of its limits.

Now, let's not forget that the actors behind the operation are assumed to be in control of the main line of defense in the DiD implementation, the RPS/ESFAS. However, even in that case this does not mean everything is lost and the cyber-physical attack will succeed. Diversity plays a fundamental role in this scenario.



One of the worst-case scenarios when implementing a digital I&C system is the Common Cause Failure (CCF or Software Common Mode Failure - SWCMF). Let's imagine that there is some kind of faulty logic in the processing modules of the digital I&C platform (TXS, Triconex, etc.) used to implement the RPS/ESFAS. The bug is extremely niche, and it will only manifest itself under very specific conditions (e.g. some floating point error), but once triggered (e.g. set-point comparison calculations) it would prevent the reactor from tripping as it is impacting all divisions/channels. A failure of the RPS to properly trip (scram) the reactor leads to a situation known as "Anticipated Transients Without Scram" (ATWS). In fact, by controlling the RPS, the attackers gain the means of forcing this accident progression, thus potentially escalating any transient into an ATWS.

In order to attenuate this scenario, NPPs tend to add to their designs additional ATWS mitigation systems such as the 'Diverse Scram System' (DSS) and the 'Diverse Actuation System' (DAS), which are implemented as a backup to the primary digital protection systems (RPS/ESFAS). These automatic systems can be based on non-computerized elements or diversified digital logic such as FPGAs (for a microcontroller-based RPS/ESFAS). Therefore, one may then assume that the cyber-physical attack will be thwarted once confronted with these systems. The truth is that this is what will most probably happen, but not necessarily in every single case.

The issue is that not every function of the primary digital protection system is duplicated into these ATWS mitigation systems ('ATWS-M' for simplicity), instead only a subset of these functions is chosen according to the safety analyses. This potentially leaves the door open for the cyber-physical attack to bypass this barrier as well. We should also note that the ATWS-M may be classified as non-safety systems (or contain both safety and non-safety elements), so they theoretically shouldn't be credited for carrying out their functions.

Another important aspect to consider is that the ATWS-M are designed to intervene only after the RPS/ESFAS has had its opportunity to do so (e.g. if a high pressure set-point in the ESFAS is set to x , in the DAS will be $x+n$), therefore there is still a certain period during which the cyber-physical attack will keep progressing.

Although some of the actuated devices can be shared between the digital primary protection system and ATWS-M systems, the design should prevent any interference between them that may impact the safety functions. However, this natural priority of the digital protection system over the ATWS-M also means that there is a chance for the attackers to attempt an attack aimed at forcing a mechanical failure in some of the important shared actuation devices, which would obviously propagate the fault to the following barriers (ATWS-M and OP).

Last, but not least, at some point after the IE (initiating event), the nuclear operators in the MCR will eventually take manual control of the situation ('OP' barrier). It is crucial to point out that trip orders manually generated from the MCR are hardwired, and are thus able to bypass the digital protection system. However, certain actuation orders are not necessarily hardwired.

The expected operator's response time can vary among countries and designs, and obviously incidents, but at least 15-30 minutes is the time required for carrying out the diagnosis of the incident and starting to perform manual actions.

For instance, let's briefly discuss how some of these systems have been implemented at Oconee NPP.

Oconee's RPS trips the reactor when detecting deviations from regular operating conditions in 10 different cases:

1. Nuclear overpower (Neutron Flux)
2. Nuclear overpower Flux/Flow/Imbalance
3. RCS (Reactor Coolant System) High Pressure
4. RCS Low Pressure
5. RCS Variable Low Pressure
6. RCS High Outlet Temperature
7. Reactor Building High Pressure
8. Loss of Both Main Feedwater Pumps
9. Main Turbine
10. Reactor Coolant Pump Power/Flux

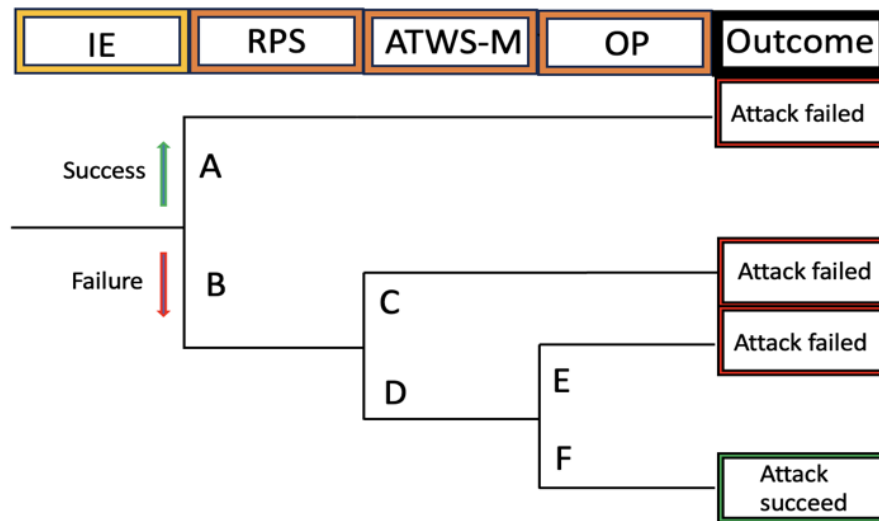
On the other hand, its 'Diverse Scram System' will automatically trip upon detecting a high RCS pressure only. However, it should be noted that a high RCS pressure is a physical symptom of many different important AOOs.

The ESFAS initiates its actuation sequences after the following conditions are sensed:

1. RCS Pressure Low
2. RCS Pressure Low Low
3. Reactor Building Pressure High
4. Reactor Building Pressure High High

Oconee's 'Diverse Actuation System' is mainly designed to mitigate Loss of Coolant Accidents (LOCA) by implementing two different systems, 'Diverse High Pressure Injection Actuation System' and 'Diverse Low Pressure Injection Actuation System'. These systems will actuate the required HPI/LPI pumps upon detecting a low RCS pressure.

So ideally, the cyber-physical attack should be able to implement the 'BDF' sequence in the following event tree.



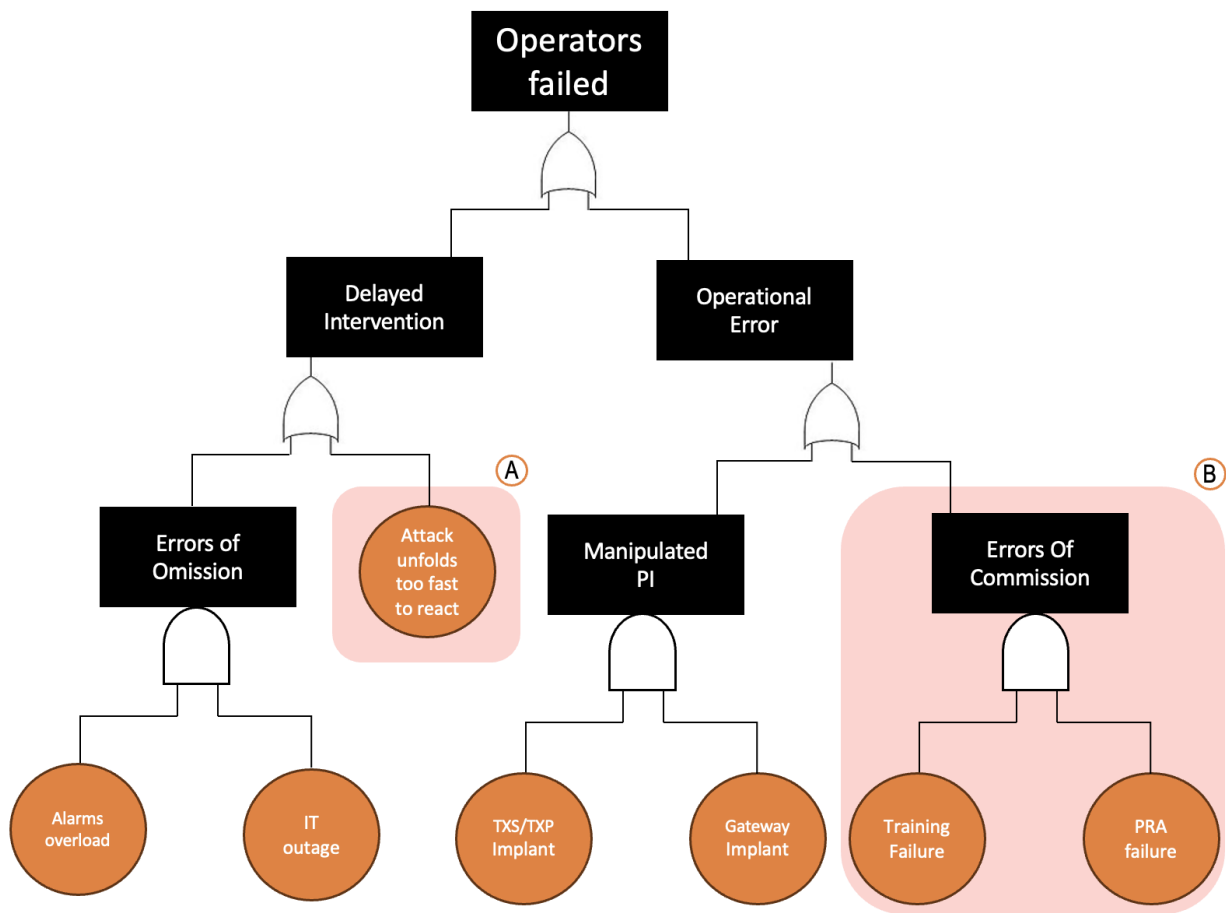
The 'Initiating Event' could be derived from modifying a set-point, or actuating a specific instrument, such as a PORV (Power Operated Relief Valve). Basically, the start of a sequence that will be triggered by the cyber-physical payload.

The first barrier (RPS/ESFAS) is assumed to be controlled by the attackers, so it would be almost certain to fail.

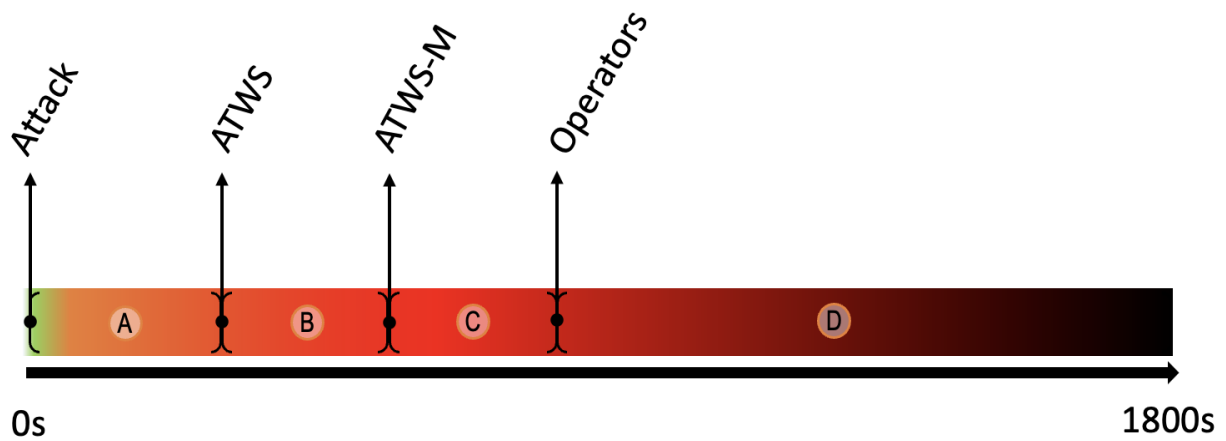
The 'ATWS-M' barrier may fail to prevent the progression of the attack in three main cases:

1. The safety analyses missed a specific sequence of events, which is being exploited by the cyber-physical attack.
2. The time interval required to reach the configured set-points is enough for the cyber-physical attack to escalate into a severe accident.
3. The cyber-physical attack has managed to induce a mechanical, or non-recoverable structural, failure into one of the shared actuation devices (e.g. by repeatedly actuating it in a short period of time).

Finally, from a deductive perspective, we can generate a sample fault tree to illustrate the idea of designing the operation in such a way that it can also induce a failure into the 'Operators' (OP) barrier. It may be too late (Delayed Intervention), the operation may have been designed to actively prevent them from properly assessing the ongoing events (a manipulated Process Information system) or the attack exploits a blind spot discovered in the safety analyses (Errors of Commission), for which the operators have not been trained. The optimal path would be either A or B, as they can almost guarantee a deterministic bypass.



To sum up the approach, let's build an ideal timeline within the first 30 minutes of the cyber-physical attack and its ideal progression. There are 4 different stages.



- A) The cyber-physical payload is activated. From this point on a sequence of physical events will occur in the nuclear reactor. Ideally, the attack should progress as soon as possible, meaning the safety posture of the plant should deteriorate rapidly. Also, if possible, this progression should cause either structural damages or not easily recoverable conditions along the way.
- B) Some of the key parameters (pressure, temperature, etc.) in the plant have reached a point that, under regular conditions, would initiate a reactor trip. However, as the attackers control the primary protection system, this escalates into an ATWS. The attack should keep progressing, thus continuously deteriorating the safety conditions over time.
- C) As the attack progresses, some of the reactor parameters should now be bad enough to lie in the scope of the ATWS-M. Even if the reactor is successfully tripped, the attack should still be progressing.
- D) At some point after the ATWS-M involvement, either successfully or not, the operators will take manual control of the situation (around 15 minutes after the IE). From the MCR they will be able to manually trip the reactor and/or issue actuation orders to mitigate the ongoing accident once they understand what is going on. At this stage, and depending on the damage sought by the attackers, there are two main options:
 - 1. The cyber-physical payload will have been designed assuming that there will be additional attacks actively trying to prevent the operators from properly understanding the situation, and acting accordingly.
 - 2. The damage is already certain regardless of the actions the operators may take.

4.4 Sample cyber-physical attack: SLOCA via Pressurizer's PSRVs

According to the premises I have previously laid out, it is time to present a theoretical-practical example of a cyber-physical attack that might potentially work in a wide representation of PWR designs. Once again, it is mandatory to refresh the disclaimer.

— Disclaimer —

It must be clear that this is a theoretical exercise, although based on realistic assumptions, technical facts, official documentation, commercial software used for severe accident simulations and reverse engineering of digital safety I&C hardware components.

I am not claiming that any of the scenarios elaborated herein may actually work in the real world, under any circumstance. I am not claiming either that this methodology represents a real threat for any of the referenced nuclear power plants or currently commissioned PWRs.

Finally, I highly discourage the use of any partial, distorted or biased interpretation of this research to create sensationalism against nuclear energy based on unscientific reasoning and FUD.

To avoid any potential misunderstanding or unfounded alarmism that might arise from just referencing specific nuclear power plants, I'll mainly use the, indefinitely suspended, US EPR design⁷⁷ as a reference to elaborate the attack. It is worth mentioning that the US EPR was conceived with a digital protection system based on the Teleperm XS platform. Let's begin.

4.4.1 - NeutronMode-IV NPP

NeutronMode-IV will be our fictional NPP where the cyber-physical attack takes place. It consists of a single unit⁷⁸, which was licensed for operating a US-EPR in baseload mode. It is located close to the 'Leidenfrost', a massive river that almost separates in two equal parts the not less fictitious Republic of Winternia, a small country submerged, over many decades, in a bitter conflict with the Kingdom of Summerniah.

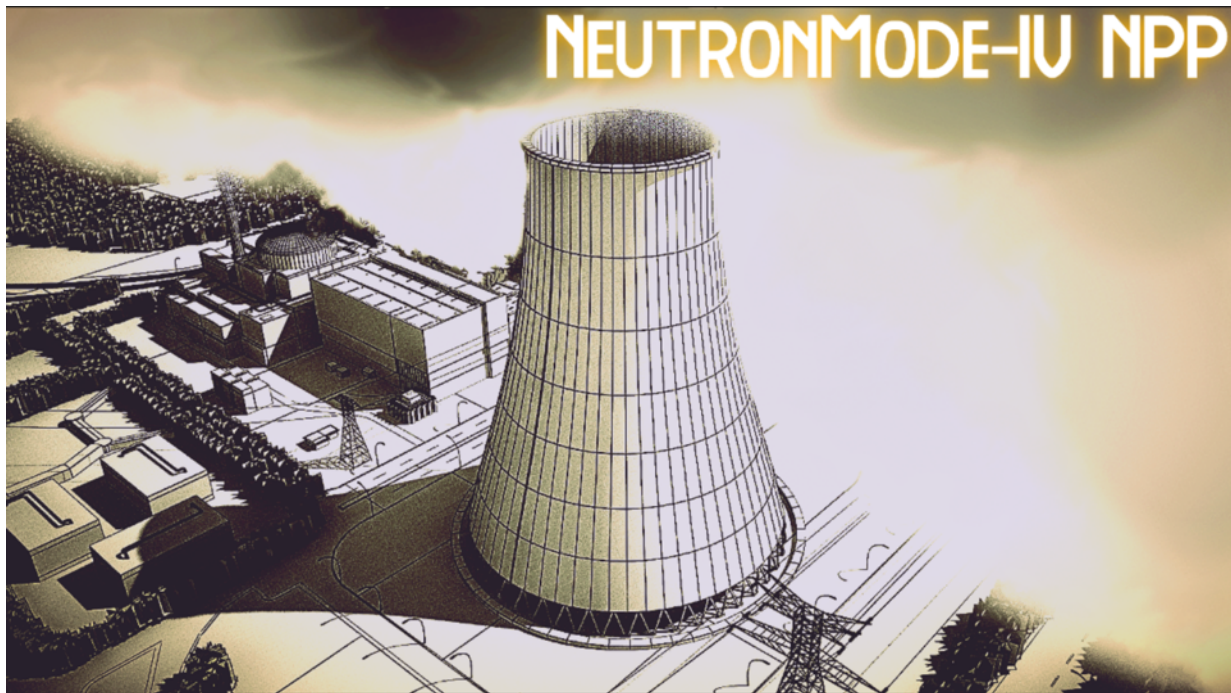
After 22 months of operating uninterruptedly since the last reload cycle, recently the plant underwent a refueling outage that lasted for several weeks. Approximately 400 additional workers, including personnel from the company that owns NeutronMode-IV but also external contractors, were working on the plant. Engineers and other technicians stayed at different hotels nearby.

⁷⁷ <https://www.nrc.gov/reactors/new-reactors/large-lwr/design-cert/epr/reports.html>

⁷⁸ <https://www.nrc.gov/docs/ML1322/ML13220A587.pdf>

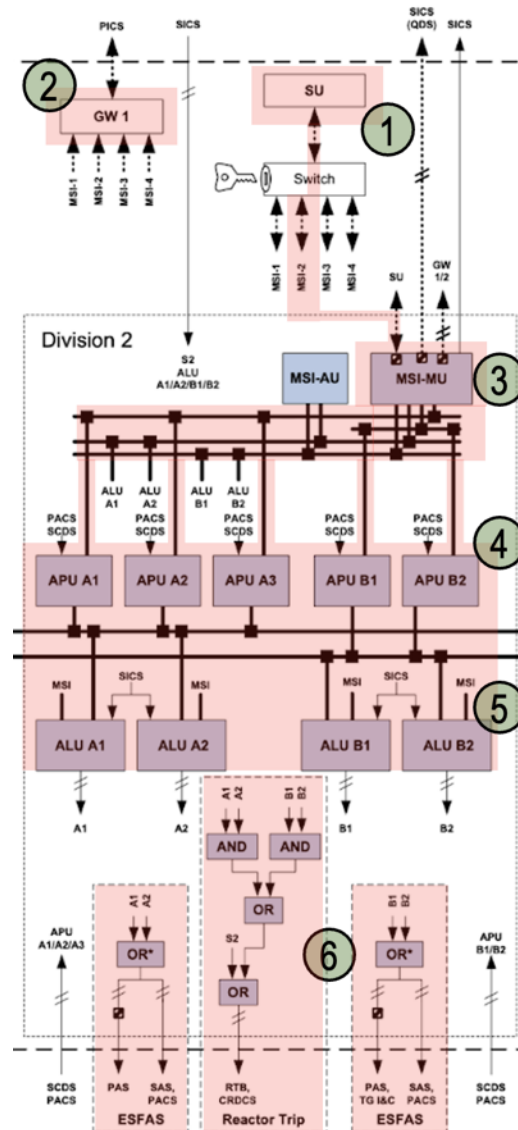
As part of the outage plan, all divisions of the digital plant protection system were updated, tested, and verified according to the plant procedures.

Some time later, on a Wednesday evening, during a particularly cold winter, a massive power cut progressively spreads across Winternia. After the first few hours of confusion, media outlets and social networks begin to be flooded with reports and rumors about an unspecified, ongoing situation at the NeutronMode-IV NPP. These messages are a mix of easily verifiable facts (NeutronMode-IV has been effectively disconnected from Winternia's national grid) and dubious reports about automatic radiation monitoring systems across the country detecting radiation spikes. Panic ensues.



This is the context for our scenario. Now let's move on to the technical details.

First, we'll elaborate what the state of NeutronMode-IV would be to enable the cyber-physical attack to be elucidated. Although the following diagram just depicts the 'Division 2' for clarity purposes, the same status applies for other divisions as well.



- 1) The attackers' implant was positioned as an 'alternative' Service Unit (either 'Portable SU' or compromised SU). During the refueling outage the malware was waiting to detect a valid connection for each of the 4 different divisions. Once this occurred, the next step was to detect the 'change operating mode' permissive (which was physically triggered by the engineers to adjust setpoints and perform tests), which marked the moment for the implant to infect the safety functional units (APUs, ALUs).
- 2) Ideally the gateway may have also been compromised. We must remember that the gateway is much more exposed than the SU, as it is continuously connected to the plant operational network. Having control over the gateway allows the attackers to control the information the operators receive, giving them the ability to modify many different plant-specific values, including alarms, and status of valves and pumps. On the other hand, operators can still maintain an accurate situational awareness since control rooms also

have physical information panels receiving hardwired data that cannot be manipulated by software implants⁷⁹.

However, a conflicting status between physical and digital information systems in the MCR can be leveraged to trick operators into performing unwanted actions, delay decisions or divert their attention from the actual underlying issue.

The following is an example from the digital upgrade at Ocone NPP⁸⁰, showing the new OPC tags for the High-Pressure Injection channels.

Page 101 of 209

CALCULATION OSC-8623, Rev. 11
RPS & ESFAS Functional Description Section 15
ESFAS Function #1
ESFAS Actuation on RCS Pressure Low

15.17 Existing Hardwired Computer Points

The existing hardwired computer points listed below will be deleted and replaced with equivalent points using computer communications (OPC gateway to OAC). New OAC point IDs and descriptions (including reset/set state messages for binary points) will be issued during detailed modification design. [Note that O1A1416 & O1A1417 remain hardwired (HW), in support of the station Core Thermal Power (CTP) calculation requirements and will also be available on the Gateway.]

Existing Point ID	Existing Description (Reset/Set state messages for binary points)	Existing Physical Range	New Destination
O1D1872	ES HP INJECTION CH A (NOT TRIPPED) (TRIPPED)	Binary	Gateway
O1D1873	ES HP INJECTION CH B (NOT TRIPPED) (TRIPPED)	Binary	Gateway
O1D1874	ES HP INJECTION CH C (NOT TRIPPED) (TRIPPED)	Binary	Gateway
O1D1875	ES HP CH A (NOT BYPASS) (BYPASS)	Binary	Gateway
O1D1876	ES HP CH B (NOT BYPASS) (BYPASS)	Binary	Gateway
O1D1877	ES HP CH C (NOT BYPASS) (BYPASS)	Binary	Gateway
O1D1890	ES CH 1 (NOT TRIPPED) (TRIPPED)	Binary	Gateway
O1D1891	ES CH 2 (NOT TRIPPED) (TRIPPED)	Binary	Gateway
O1A1416	RC LOOP A WR PRESS 1	0 to 2500 psig*	Gateway/HW
O1A1417	RC LOOP B WR PRESS 1	0 to 2500 psig*	Gateway/HW
O1A1418	RC LOOP A WR PRESS 2	0 to 2500 psig	Gateway

* The hardwired input signal to the OAC is 0 to 10 VDC, representing 0 to 2500 psig.

- 3) Please note that there is no need to compromise the MSI. The MSI (main unit), by design, enables a logical bi-directional communication between the SU's implant and the safety network. The MSI is designed to route (expected) messages, coming from the SU to the specific functional unit. Each SVEx/SCPx in the APU/ALU is identified by an ID which is used as a destination field in the internal network messages.
- 4) The APUs were plausibly compromised. As we discussed, the approach may be different depending on the requirements of the cyber-physical attack. However, it is completely feasible to assume that sophisticated attackers were able to gain persistence (either RAM-only or reset-resistant) by deploying an implant while the unit was in 'DIAGNOSTIC' mode.

⁷⁹ <https://www.nrc.gov/docs/ML1322/ML13220A740.pdf>

⁸⁰ <https://www.nrc.gov/docs/ML0910/ML091050333.pdf>

- 5) The scenario elaborated in the previous point also applies for the ALUs. However, there is something important to discuss in this specific scenario.

In the previous 'Implementation' section I discussed the various alternatives used to prevent the operators from manually intervening during the ongoing cyber-physical attack. In the US EPR design, there are different levels of manual controls, some of them bypass the PS (a TXS-based RPS/ESFAS), but others do not. This is important, because it enables the implant to not only control the automatic actions of the PS but also influence a significant number of manual actions that the operators are credited to perform from the Main Control Room.

It is important to remember that the ESFAS is required to actuate at every possible mode of operation (shutdown, full power, etc.).

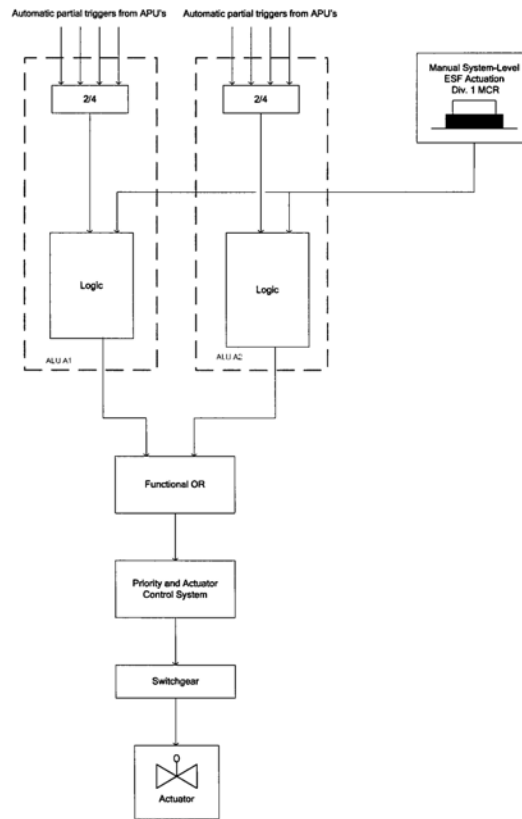
Manual Controls that do not bypass the PS, and therefore are under malware's control.

1. System-level ESFAS actuation sequences originating from the SICS.
The following 18 system-level manual actuation sequences are routed through the PS, thus being subject to the control of the ALU's implant.

- Safety Injection System Actuation
- Emergency Feedwater System Actuation
- Emergency Feedwater System Isolation
- Partial Cooldown Actuation
- Main Steam Relief Isolation Valve Opening
- Main Steam Relief Train Isolation
- Main Steam Isolation
- Main Feedwater Isolation
- Containment Isolation
- Chemical and Volume Control System (CVCS) Charging Isolation
- CVCS Isolation for Anti-Dilution
- Emergency Diesel Generator (EDG) Actuation
- Pressurizer Safety Relief Valve Opening (Brittle Fracture Protection)
- Steam Generator Isolation
- Reactor Coolant Pump Trip
- Main Control Room Air Conditioning System Isolation and Filtering
- Turbine Trip on Reactor Trip Initiation
- Hydrogen Mixing Dampers Opening

In the following image (US EPR) we can see the functional implementation to trigger these manual system-level actuation sequences.

Figure 8-3—Manual System-Level ESFAS Actuation Sequence (One Division)



For instance, for certain events, such as a Steam Generator Tube Rupture (SGTR), the safety procedures establish that the operator is credited to perform a manual system-level initiation.

Manual Controls that bypass the PS, and therefore cannot be controlled by the malware:

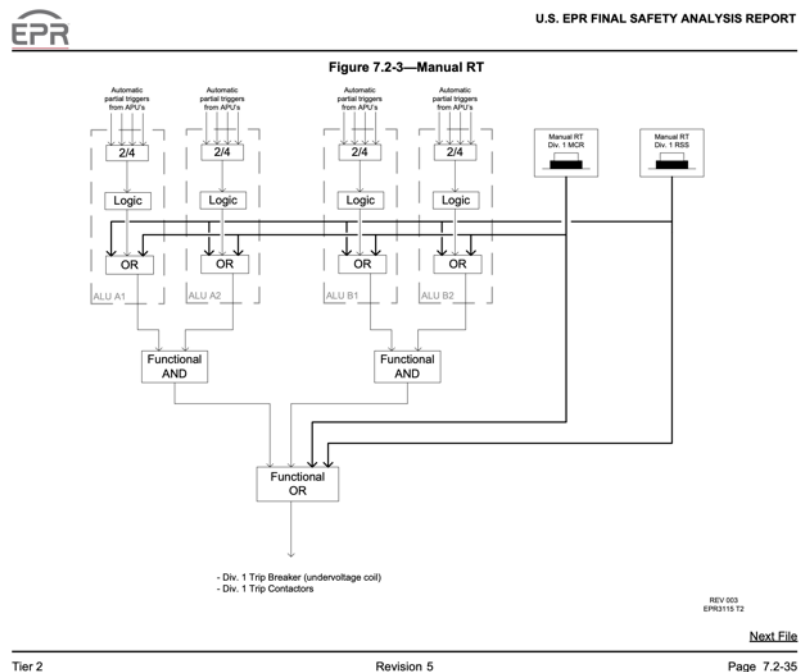
1. Predefined or component-level actuation orders originating from PICS and SICS that are directly routed (hardwired) to the Priority Actuation Control System (PACS), or implemented in the DAS⁸¹.
 - Reactor Trip (SICS/DAS/PACS)
 - EDG start (SICS/PACS)
 - Component controls to support diesel generator loading (emergency diesel generators or SBOs) (SICS/PACS)
 - EFW actuation (SICS/DAS/PACS)
 - Operation of EFW for long-term SG level control (SICS/PACS)
 - SI switchover to hot leg injection (SICS/PACS)
 - MSIV closure (SICS/PACS)
 - Feedwater isolation (MFW and EFW) (SICS/PACS)

⁸¹ <https://www.nrc.gov/docs/ML1307/ML13073A605.pdf>

- Initiation of medium head safety injection (MHSI) (SICS/DAS/PACS)
- Control of MHSI (SICS/PACS)
- Extend partial cooldown (SICS/PACS)
- Depressurize RCS with pressurizer sprays (SICS/PACS)
- Actuation of extra borating system (EBS) (SICS/PACS)
- Control room HVAC reconfiguration (SICS/PACS)
- CVCS isolation (SICS/PACS)
- MSRT control (SICS/PACS)
- Stage 1 containment isolation (SICS/DAS/PACS)
- Opening of containment hydrogen mixing dampers (SICS/DAS/PACS)

2. Manual Reactor Trip orders.

As the following logic shows, as opposed to the previous ESFAS logic, the manual Reactor Trip orders are hardwired both to the ALUs and the hardware-based 'Functional OR', thus ensuring an independent path to the Reactor Trip breakers, even when the RPS has been compromised. It should be noted that the main idea behind this design is not specifically to prevent cyber-attacks but CCF/SWCMF (Common Cause Faults/Software Common-Mode Faults) scenarios. It may be sooner or later, but it must be assumed that the operators will manage to successfully trip the reactor, in any given case.



6) Essentially, automatic actuation sequences and reactor trip orders can be plausibly placed under malware's control.

4.4.2 - “Cyber Three Mile Island”

In certain professional contexts there is a taste for creating analogies between some notorious tragic events in history and certain scenarios that may have a ‘cyber’ component. As a result, reading stories, articles or press releases talking about “Cyber Pearl Harbor” or “Cyber 9/11” is not uncommon. One may think that this research would be a perfect occasion to forge the “Cyber Chernobyl” parallelism, but even if I liked that kind of terminology, which is not the case, I wouldn’t ever choose it. Instead, it would in fact be much more accurate and realistic to talk about “Cyber Three Mile Island”. If we ever see a successful cyber-physical attack against a modern PWR, the worst-case outcome would likely be closer, both in progression and limited impact, to ‘Three Mile Island’ than to ‘Chernobyl’ or ‘Fukushima’.

There are many reasons for this, but above all Chernobyl was a very specific RBMK design, with a positive void coefficient and no containment, totally the opposite of what we have in a PWR. Fukushima was an accident mainly driven by extraordinary external events. On the other hand, some of the underlying issues behind the sequence of events that led to the Three Mile Island accident, are actually very close to being ‘exploitation primitives’ for a cyber-physical attack against a PWR. This will be developed further in the next sections, but first of all let’s refresh a couple of basic concepts.

1. Cooling the fuel is one of the three main safety principles (the three ‘C’s) for NPPs.
2. The ‘Reactor Coolant System’ is a closed-loop circuit designed, among other things, to remove the heat from the fuel.

The main idea is obvious: if there is no coolant, you cannot cool the fuel. If the fuel is not cooled, then it is going to melt.

This scenario is known as a ‘Loss of Coolant Accident’ (LOCA). Depending on the flow of coolant being leaked, this kind of accident is usually divided into Small LOCA (e.g. a leak in a pump) and Large LOCA (e.g. a huge crack in a hot leg).

There is only one commercial PWR that has suffered a severe accident where a (Small) LOCA played a fundamental role: Three Mile Island. This particular accident changed many things about the way that safety was approached for nuclear designs, so lessons were definitely learned.

An attack on an NPP is a one-shot attempt, if you fail there is not going to be a second chance in a long time, if ever again. Therefore, it’s not a bad idea to take an approach similar to what has been found to work in a real-world severe accident.

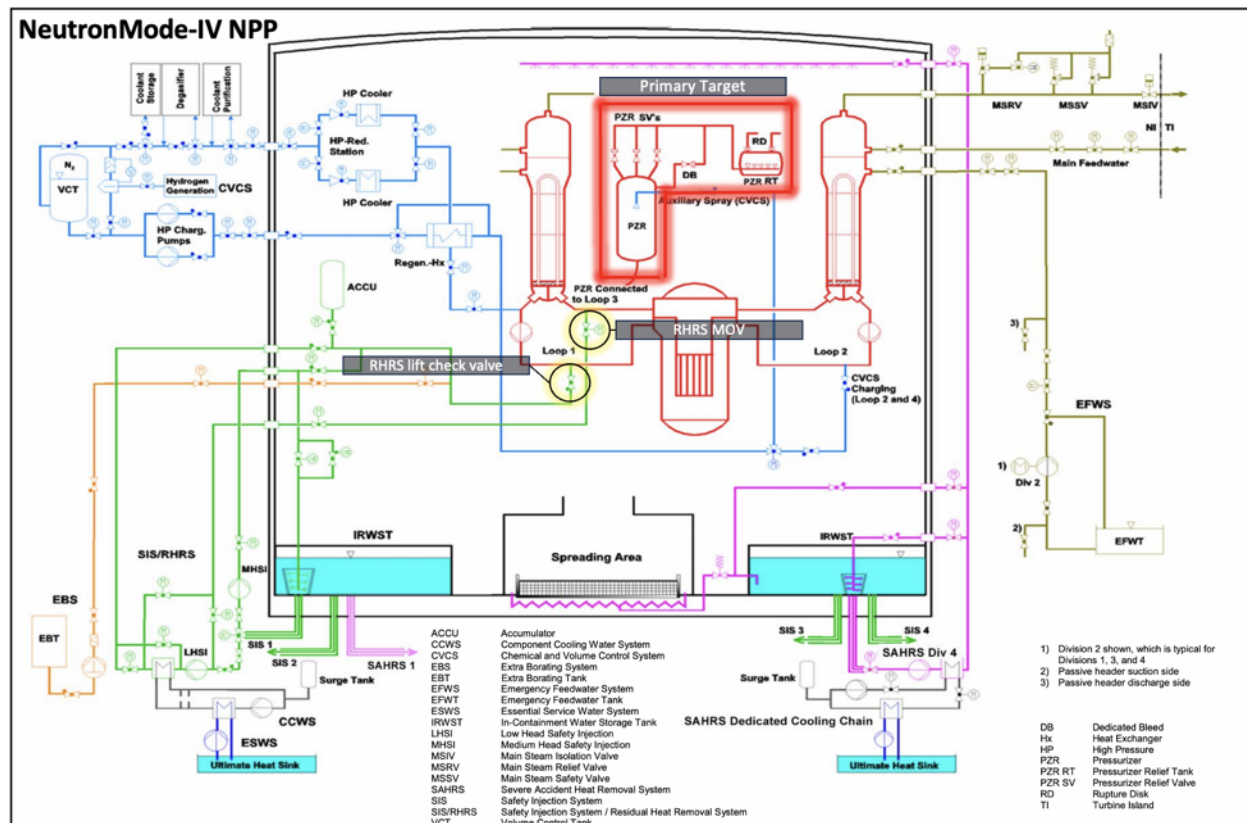
4.4.3 - No pressure

When you are trying to turn a specific software vulnerability into a working exploit, you need to leverage every possible technique to influence the target in your favor. This means finding exploitation primitives that allow you to perform some basic operations required for a successful exploit, such as leaking a memory address, or an arbitrary memory write.

In the same line, when your target is a nuclear reactor, and you are looking to influence key reactor parameters (pressure, temperature, etc.) you also need to find the right 'exploitation' primitives, because you can't just 'increase' pressure here or 'decrease' temperature there.

So, let's say that the required objective is to cause a, relatively fast, decrease in the coolant inventory of the Reactor Coolant System. What are the options?

First of all, let's take a look at the fluids architecture of the US EPR.



Almost all of these systems are potentially prone to enable a LOCA scenario if somehow the attackers manage to cause damage in their associated piping. Although, this is far from being an easy task, and even if it is eventually achieved, there are still a bunch of different safety (and non-safety) systems that can work to mitigate a LOCA: the Safety Injection System (SIS) with its Medium Head Safety Injection (MHSI) and Low Head Safety Injection (LHSI), Accumulators

(passive), the Residual Heat Removal System (RHRS), the Extra Borating System (EBS), the In-Containment Refueling Water Storage Tank, or the Chemical and Volume Control System (CVCS).

A possible approach would be to target the isolation points between those systems designed to work in different plant operations modes. This means that they are expected to operate within a specific range of pressure and temperature⁸² mainly, so by taking them out of their working regime we can expose these structures and components to thermal shocks or overpressure. As we have seen, a notable example of this approach was Stuxnet.

So, we can see, the Residual Heat Removal System (RHRS) is designed to keep the fuel cooled during shutdowns by removing the decay heat still present in the core (~6%). Thus, its Motor-operated Valve (RHRS MOV in the image above), which is connected to the hot leg, is interlocked with a permissive that prevents its inadvertent opening while the plant is operating at full power⁸³.

For shutdown cooling operations, each SIS/RHRS train is aligned to take suction from the corresponding RCS hot leg, pump the hot reactor coolant through the corresponding LHSI HX where it is cooled by the corresponding CCWS train, and return the reactor coolant to the RCS through the corresponding cold leg. Interlocks (permissive P14, refer to Section 7.2.1.3.9) prevent alignment of the SIS/RHRS to the RCS while RCS pressure and temperature is greater than approximately 464 psia and 350°F. This feature protects the SIS/RHRS components from overpressure due to exposure to the RCS pressure during reactor operation (intersystem LOCA). Additional features addressing intersystem LOCA are discussed in Section 5.4.7.2.2.

If that permissive could be bypassed, the malicious actuation order would open the valve, thus breaking the isolation between the RHRS and the RCS at full power. It is worth mentioning that interlocks and permissives are under the control of the PS⁸⁴, so when ALUs are compromised and permissives are exclusively verified by its application software, there is nothing to prevent the required set of ALUs (usually 2-out-of-4) in the compromised ESFAS from cooperating to maliciously actuate the target field devices.

The outcome is that the RHRS will receive a brutal pressure and temperature discharge coming from the hot leg, which might plausibly induce a break in the RHRS piping, thus enabling a LOCA. However, that scenario is not certain: ideal safety engineering states that if a system A (low pressure) is not completely isolated from a system B (high pressure), materials in A should be able to withstand⁸⁵ the highest pressure.

⁸² <https://www.nrc.gov/docs/ML1218/ML12187A024.pdf>

⁸³ <https://www.nrc.gov/docs/ML1322/ML13220A690.pdf>

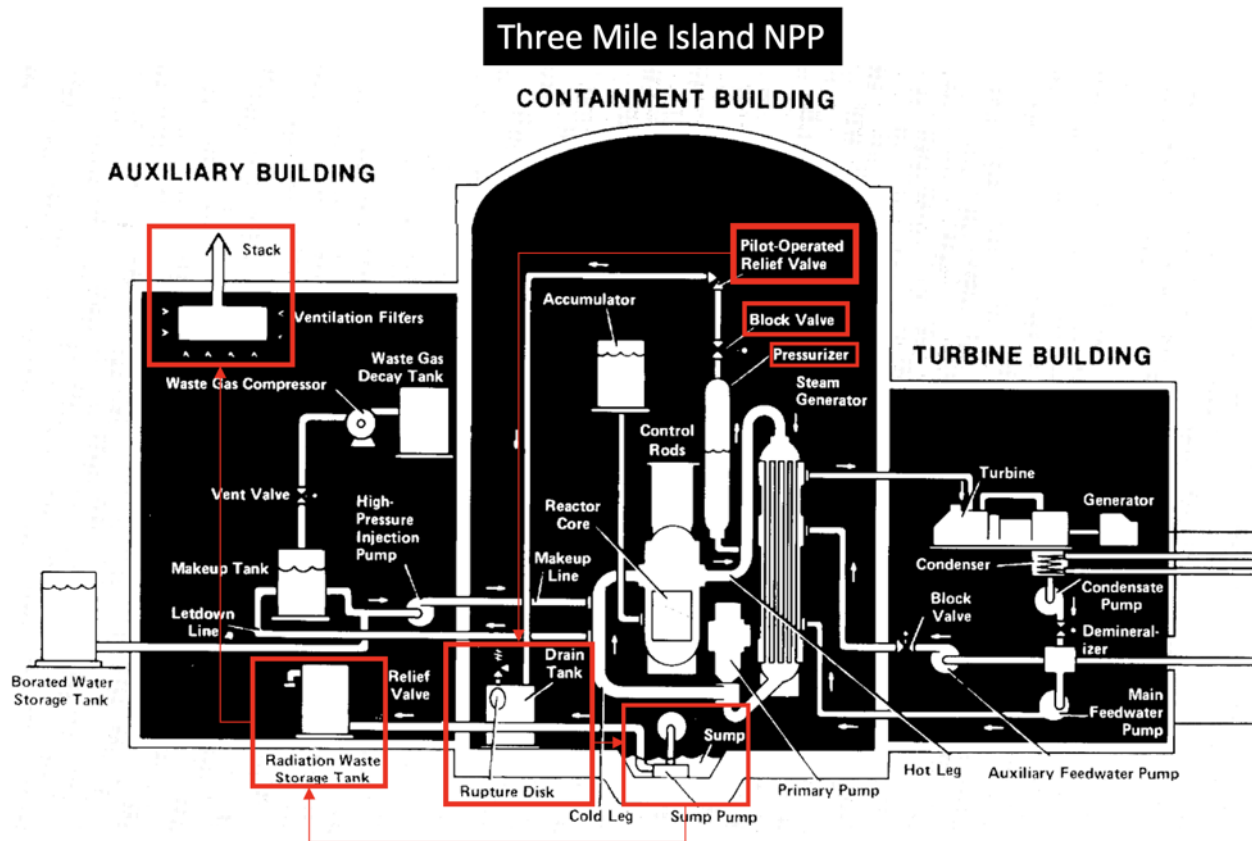
⁸⁴ <https://www.nrc.gov/docs/ML1322/ML13220A721.pdf>

⁸⁵ <https://www.nrc.gov/docs/ML1322/ML13220A642.pdf>

Therefore, it seems a better plan to go after a more deterministic approach: the Pressurizer's Pressure Safety Relief Valves.

4.4.4 - Initiating event

Before elaborating the rationale behind choosing the pressurizer as the primary target, let's briefly analyze the sequence of events during the first 4 minutes of the Three Mile Island accident.



1. 0s - 4:00:37 AM

Feedwater pumps trip so the turbine is automatically tripped too, because the steam demand cannot be fulfilled. The auxiliary feedwater pumps activate to make up for the lack of main feedwater, but it turns out that the valves were closed due to a botched maintenance operation. Without having enough water in the Steam Generators, the heat generated in the core is not being transferred to the secondary circuit, so the pressure in the primary circuit increases.

2. + 6s - 4:00:43 AM

The Pilot-Operated Relief Valve (PORV) in the Pressurizer opens after reaching the high pressure setpoint. This valve is directly connected to the Reactor Coolant Drain Tank (RCDT), where the vented coolant is being deposited.

3. + 10s - 4:00:47 AM

The RPS triggers an automatic Reactor Trip after reaching the RCS High Pressure setpoint. From now on, the amount of heat generated in the core is drastically reduced as the chain reaction has been stopped. The overall pressure in the RCS decreases as the reactor is basically just producing decay heat.

4. + 16s - 4:00:53 AM

The PORV that was still venting coolant is expected to close as pressure is now decreasing after the shutdown. However, it failed and failed open. Without implementing a check-back signal to confirm the status of the valve, the operators merely received the information that a 'close' actuation order had been issued, therefore assuming it had succeeded.

5. +120s - 4:02:37

The stuck-open PORV keeps venting coolant from the primary circuit to the RCDT, basically creating a LOCA scenario. Pressure keeps decreasing in the RCS, enough for the High-Pressure Injection System to be initiated, which tries to mitigate the loss of coolant by injecting water from the makeup tank.

6. +240s - 04:04:37

The RCDT is not prepared to receive a continuous flow from the PORV, so its 'Rupture Disk' bursts, thus leaking the primary coolant, which is radioactive as it is in direct contact with the core, to the containment building. This radioactive water is collected by the sump pump and pumped to the storage tank in the auxiliary building. At some point, this tank also overflows, thus letting the radioactive water escape into the environment via the building vent stack.

During the next hours, the situation continued to deteriorate, the LOCA progressed to a point where the core was left uncovered and partially melted. What happened in those 4 minutes fits very well with the generic characterization of our cyber-physical attack.

Essentially in 4 minutes, the 4 D's have been fulfilled:

a) Disaster

Although still a minor release, the slightly radioactive coolant that has been vented to the atmosphere paves the way for it to be considered a 'Disaster'.

b) Destruction

At least one component inside the Reactor Building has been physically damaged, since the RCDT's rupture disk burst.

c) Disruption

A short-term disruption is ensured due to the Reactor trip. On top of that, we have physical damages in the RCDT which ensure, at least, a medium-term disruption.

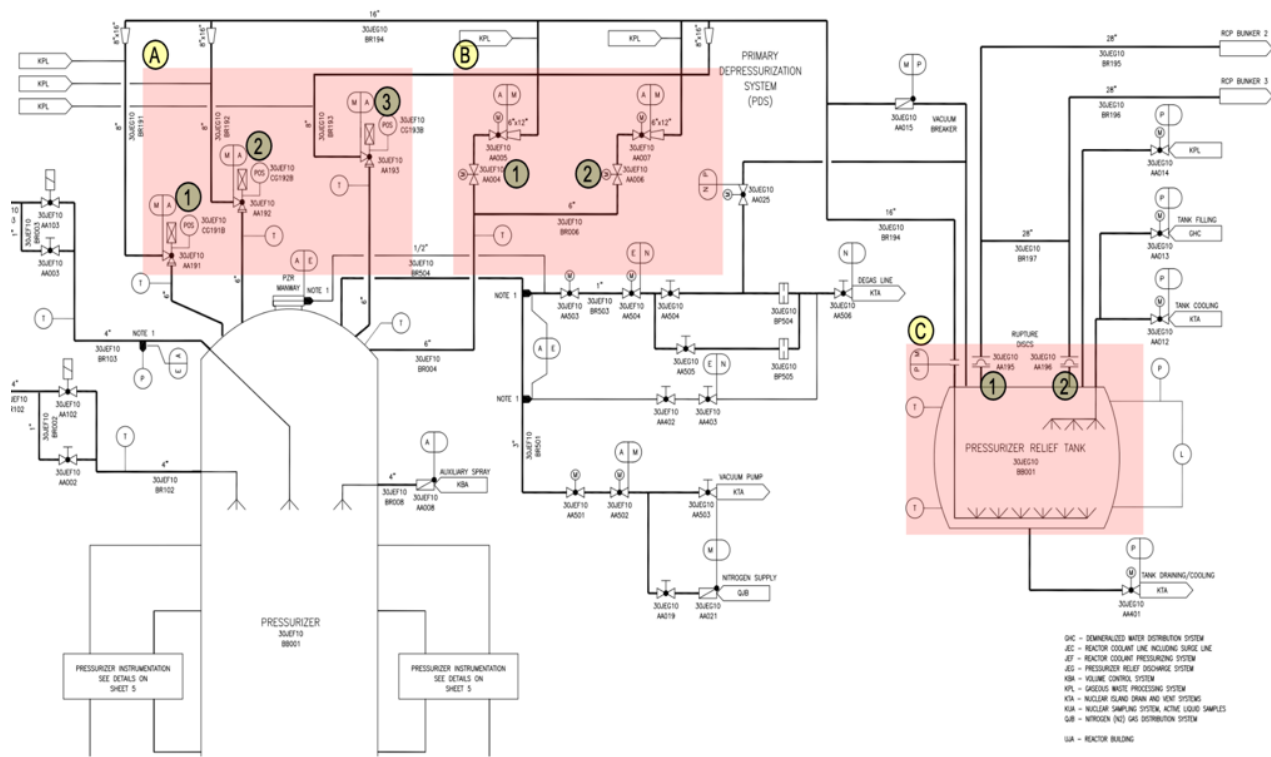
d) Deception

Relying on the previous elements it is possible to launch a disinformation campaign that can mix facts, half-truths and just plain lies. This can be accompanied with more traditional cyber-attacks targeting radiation monitoring systems to fabricate radiation levels, thus depicting a much more 'dramatic' situation than it actually is.

Now the question is: is it possible to transform this Three Mile Island initial sequence into a cyber-physical payload? Let's elaborate on it.

Pressure Safety Relief Valves

As we can see in the following image (US EPR), the design of the pressurizer in the US EPR includes three Pressure Safety Relief Valves (PSRV) (Block A. 1, 2, 3), which are connected to the Pressurizer Relief Tank (PRT). This tank in turn, has two pressure relief devices in the form of two rupture disks (Block C, 1 and 2). At a first glance, it may look pretty similar to Three Mile Island architecture, but obviously things have changed since then, so there are similarities but also very important differences. Interestingly, as they are both safety and relief valves, these PSRVs don't have any block valve that could isolate them, as opposed to the two MOVs (Block B, 1,2) in the Primary Depressurization System (PDS).



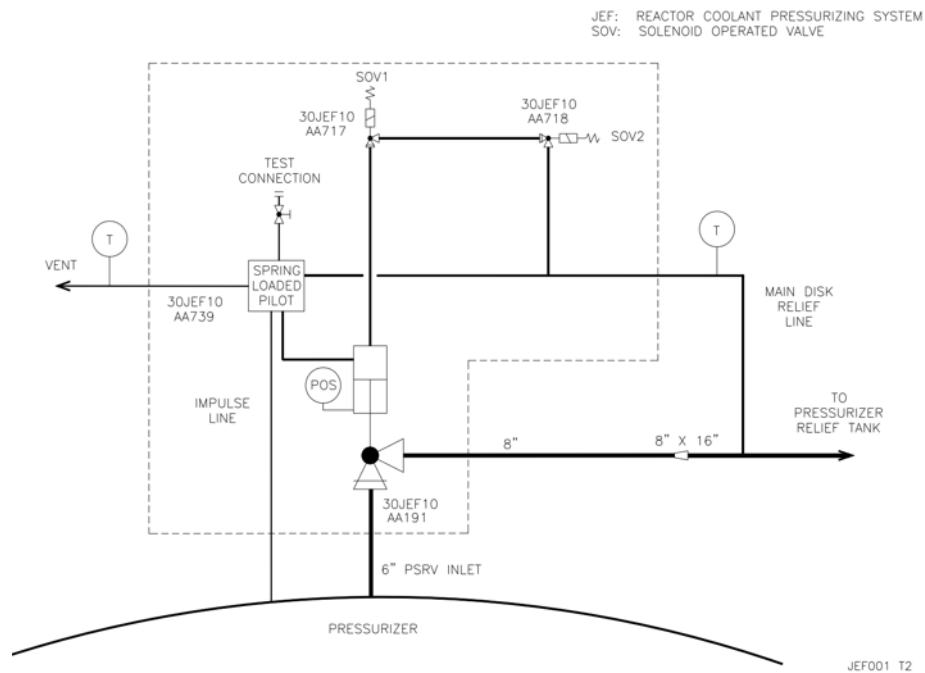
These PSRVs have a dual design⁸⁶ according to the operation mode of the plant:

- Power Operation

The PSRVs are considered passive devices. Each PSRV contains a spring-operated pilot valve that will passively actuate the valve, when the mechanically configured setpoint is reached.

- Cooldown/Low temperature

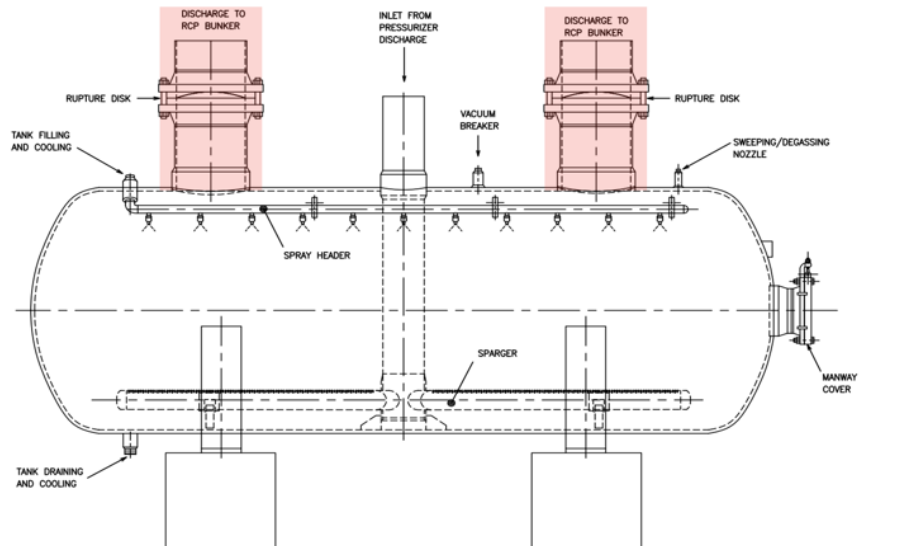
2 of the PSRVs are considered active devices, which can be either automatically actuated by the ESFAS (mainly for the Low Temperature Overpressure Protection sequence, LTOP) or manually by the operators from the MCR. This latter procedure requires manually validating a LTOP permissive (P17). Each of the PSRVs contains two solenoid-operated valves mounted in series to prevent spurious actuations.



Pressurizer Relief Tank

The Pressurizer Relief Tank (PRT) receives the discharge from the 3 PSRVs and the 2 MOVs of the PDS. Once the pressure inside the tank reaches a specific setpoint, its rupture disks will burst. The discharged coolant will then be routed through a safe path, to avoid any impact on other systems, towards the containment sumps, where it will be reused by the Safety Injection System, thus enabling its recirculation.

⁸⁶ <https://www.nrc.gov/docs/ML1322/ML13220A690.pdf>



According to these elements we can now conclude that it would be plausible to partially implement the initial Three Mile Island sequence in a cyber-physical payload:

- It is possible for a compromised ESFAS to actuate 2 of the 3 PSRVs. This would be comparable to the stuck-open PORV in Three Mile Island (TMI).
- A continuous discharge of the PSRVs will eventually cause the rupture disks in the PRT to burst. However, the discharged coolant will be collected and reused as part of the Safety Injection System. Therefore, no release of radioactivity will be produced. This is different from TMI.
- The position of the PSRVs is indicated in the MCR. While a compromised gateway can potentially modify these indications in the PICS, hardwired indications cannot be modified. Therefore, operators will know that the PSRVs are open. This is different from TMI. Valves are usually equipped with acoustic and temperature detectors so as not to mistake its actual status.
- The Reactor will be tripped due to low pressure detected in the primary circuit.

Therefore, with regards to the 4 D's we have:

Disaster	✗
Destruction	✓
Disruption	✓
Deception	✓

Is there any chance of reaching the ‘disaster’ level with this initiating event? There is something in the US EPR design that may actually enable this outcome.

4.4.5 - A matter of priorities

The US EPR PRA does not contemplate this scenario, but this is expected because PRAs do not model events assuming a compromised PS (there is no way to quantify events according to a cyber-attack or terrorism), but mainly CCF/SWCFM.

AREVA NP Inc.

U.S. EPR Probabilistic Risk Assessment
Methods Report

ANP-10274NP
Revision 0

Page 2-39

Table 2-1—Example Table of Initiating Events Selection for at Power

NUREG/CR-5750 Initiating Events	US EPR Initiating Events
Loss-of-Coolant Accident (LOCA)	
Large Pipe Break LOCA	LLOCA
Medium Pipe Break LOCA	MLOCA
Small Pipe Break LOCA	SLOCA
Very Small LOCA/Leak	Not modeled. Assumed that normal charging will maintain inventory.
Stuck Open: Pressurizer PORV	Not applicable
Stuck Open: 1 Safety/Relief Valve	Design makes this highly unlikely. Included in SLOCA.
Stuck Open: 2 Safety/Relief Valves	Not Modeled
Reactor Coolant Pump Seal LOCA	RCP seal LOCAs are evaluated within event trees.

However, a pretty similar issue is part of the US EPR transient and accident analyses⁸⁷: Inadvertent Opening of a Pressurizer Safety Relief Valve (IOPSRV).

The IOPSRV causes a loss of reactor coolant inventory that cannot be offset by the chemical and volume control system (CVCS). This condition causes primary system depressurization and a decrease in reactor coolant density. In the early phase of the event, the reactor power is determined by reactivity feedback (moderator density) and the reaction by the rod position controller (automatic rod control system).

The IOPSRV is not exactly the same situation, as a spurious actuation is assumed. Anyway, both cases would be modeled, and handled, in the same way, as a SLOCA. However, from the operators' perspective, there is a big difference between addressing a SLOCA due to a spurious actuation and trying to do so when there is a malicious PS actively working to prevent it.

So, let's recap the situation. As a result of the cyber-physical attack, we have that:

1. 2 PRSVs are actuated.
2. Automatic and manual System-Level ESFAS actuations designed to deal with a SLOCA can be inhibited. As we have seen, these actuation sequences cannot bypass the PS, and therefore are fully controlled by the compromised ALUs.

⁸⁷ <https://www.nrc.gov/docs/ML1322/ML13220A933.pdf>

3. The operators are then required to handle the situation via manual controls. However, they only can use those controls that can bypass the PS, otherwise the compromised ALUs will block their actuation orders.

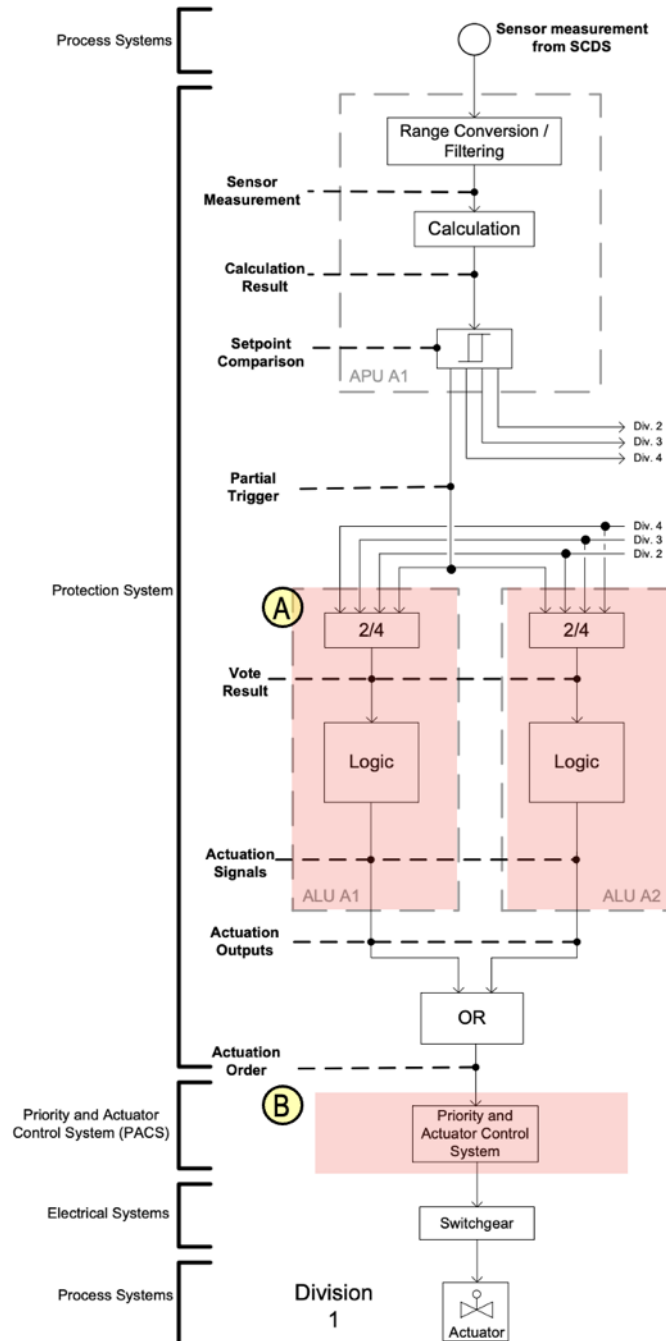
There is, however, still a system that can actively deny the operators' manual actions even assuming that they are sending actuation orders that bypass the PS. The crucial system in this scenario is the Priority and Actuator Control System (PACS).

The PACS performs prioritization of signals received from different I&C systems, and manual controls. The TXS AV42, a priority actuation and control (non-programmable) module⁸⁸ is part of the core implementation for this system.

The diagram on the right (US EPR) represents the actuation logic, showing:

- A) The compromised ALUs, that can block any actuation order that is routed through the PS.
- B) The PACS system, which is independent of the PS.

The following paragraph⁸⁹ introduces the issue.



⁸⁸ <https://www.nrc.gov/docs/ML0633/ML063380086.pdf>

⁸⁹ <https://www.nrc.gov/docs/ML1322/ML13220A721.pdf>

The U.S. EPR I&C design allows for multiple I&C systems to send requests to a given actuator. To make certain that each individual actuator executes the proper action for the given plant condition, priority management rules for the PACS are provided. The following systems inputs to the PACS are listed in order of priority:

- PS/DAS.
- SAS.
- SICS.
- PAS.

The top priority in the PACS is for the PS and DAS system, which results in an interesting scenario. This means that actuation orders from the PS/DAS will have priority over any other order received from the SAS (Safety Automation System), SICS (Safety Information and Control System, manual controls in the MCR) and PAS (Process Automation System).

As a result, the operators may be trying to manually close a valve via the SICS, but at the same time the compromised PS is continuously sending actuation orders to keep it open. As the PS has priority, the operator's orders are expected to fail.

The only chance for operators to match the PS priority is to use those actuation orders implemented through the DAS. However, as it has been previously shown, these are limited to the following:

- Stage 1 containment isolation (SICS/DAS/PACS)
- EFW actuation (SICS/DAS/PACS)
- Reactor Trip (SICS/DAS/PACS)
- Initiation of medium head safety injection (MHSI) (SICS/DAS/PACS)
- Opening of containment hydrogen mixing dampers (SICS/DAS/PACS)

Fortunately for the operators, the MHSI is one of the actuation orders implemented in the DAS. The MHSI system represents a key mechanism in increasing the coolant inventory during a SLOCA scenario.

However, even if the operators manage to win the 'race condition' by using the DAS, the malicious PS can actuate the target component again, leading to the same scenario over and over. On top of that, actuating a device many times is far from being recommended, as the probability of a mechanical failure increases.

Therefore, the ability of the cyber-physical attack to escalate the SLOCA (or any other) initiating event into a 'Disaster' will depend on two main factors:

- 1) The time required for the cyber-physical attack to force a non-return point in the safety conditions of the plant.

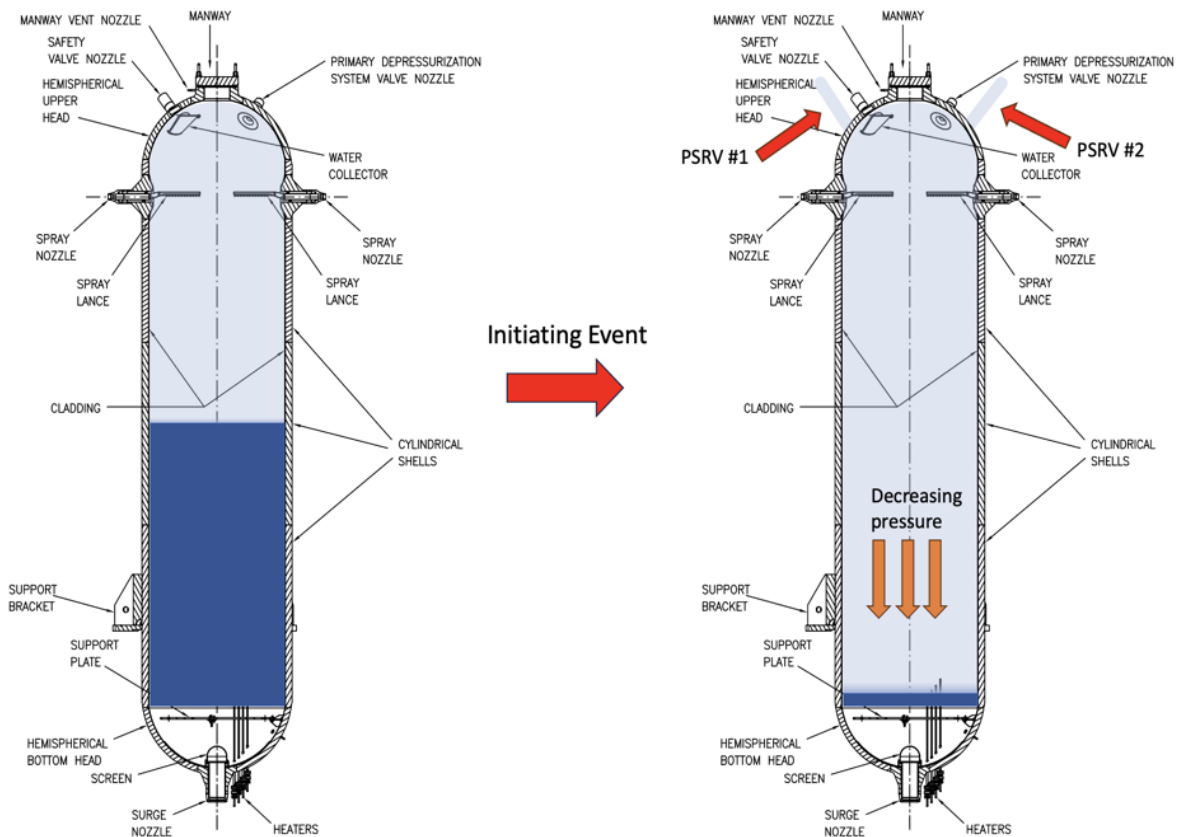
- 2) The capacity of the operators to deal with an 'adversarial PS'. A very specific situation for which, a priori, they may not have been trained.

4.4.6 - Escalating a SLOCA into a severe accident

In order to understand how a SLOCA can be turned into a severe accident, we'll need to dust off some basic physics. Also, we'll be able to apply some of the concepts elaborated in the "Physics of Pressurized Water Reactors" section. Let's start.

The pressurizer is in charge of keeping the RCS pressure under the expected operating limits by maintaining a saturated mixture of water and steam at equilibrium. Let's analyze what happens when the 2 PSRVs on top of the pressurizer are actuated by the compromised ALUs.

The steam 'bubble' that regulates the pressure in the pressurizer immediately begins to be vented. This means less atoms in the same volume so there will be fewer particles 'hitting' the water, thus decreasing the pressure. As soon as this is sensed by the pressurizer control system, the heaters start to boil water to make up for the loss of pressure. However, the PSRVs keep venting the generated steam so there is no way the heaters can fulfill their function. Actually, at this point the heaters are basically working in favor of decreasing the water inventory in the pressurizer. When the water level is too low, the heaters stop working to avoid being damaged. The pressure keeps decreasing.

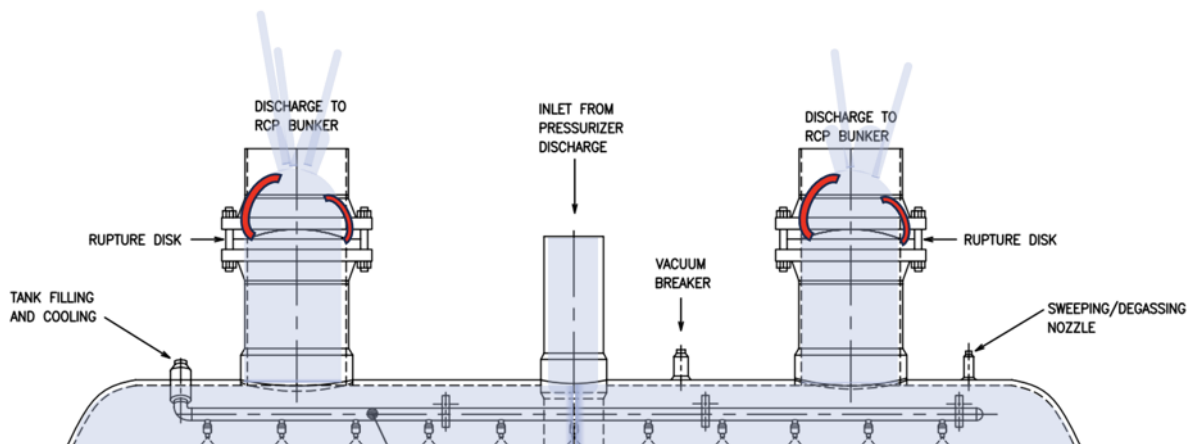


The primary circuit is a closed-loop fluid system, so although pressure is not the same for all its elements (for instance, due to its elevation) any pressure variation will equally impact them. Therefore, a decreasing pressure in the pressurizer means that the pressure in the RCS is also decreasing.

When the RCS Low Pressure setpoint is reached (~ 125 bar), the Reactor will be tripped by the PS. Right, we have an 'adversarial PS', but it doesn't really matter whether it tries to prevent this trip from happening or not. At most, a few seconds for the progression of the initiating event can be gained, but eventually the DAS would trip the reactor anyway.

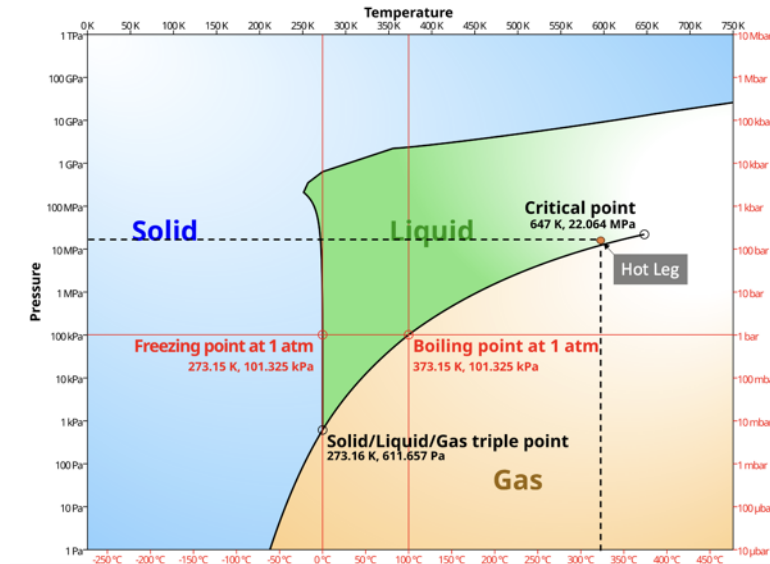
If the reactor is tripped, the turbine cannot continue operating either because there would be a load imbalance, as its demand of steam cannot be provided any more. Therefore, the turbine will be automatically tripped as well. However, the Steam Generators (SGs) are still working, so the steam is now transferred to the condenser, which is outside the nuclear island. It would be possible to force the SGs to dump steam into the environment directly instead of using the condenser, but let's not forget that coolant flowing through the secondary circuit is not radioactive (unless isolation between primary and secondary circuits is broken, for instance if some of the U-Tubes break). It is true that for deception purposes, it would serve as a visual 'proof' of the cyber-attack, which then might be subject to misinterpretations. Apart from that, and the fact that it would be quite noisy, it doesn't represent any radiological threat.

Some minutes after the reactor has been tripped, the rupture disks in PRT will burst due to the discharge flow coming from the 2 stuck-open PSRVs.



At this point, operators will be receiving many different alarms, including the loss of subcooling margin. The coolant circulating in the RCS is at high pressure to prevent it from boiling. The pressurizer was keeping a constant pressure of ~ 155 bar and the hot leg temperature should be around 330°C during power operation. However, the pressure is now lower so if the temperature is still high enough (see the following phase diagram of water), coolant in the core will boil.

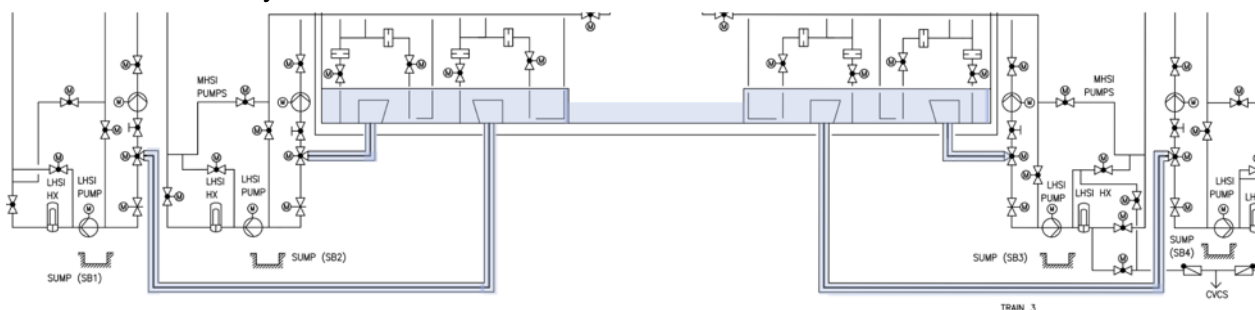
And that's really bad.



Before that point is reached, in a normal scenario the ESFAS would initiate those actuation sequences required to restore the coolant inventory that is being lost, and to keep removing decay heat from the core. Later, the operators would intervene to perform the required manual operations they are credited with.

However, in this case, the adversarial ESFAS will explicitly prevent that from happening. Therefore, the compromised ALUs will inhibit any attempt to initiate both the Safety Injection System and the Residual Heat Removal System. On top of that, we should bear in mind that the 2 PSRVs cannot be isolated because they do not have a block valve downstream, and any operator's attempt to close them manually from the control room, will be subsequently reverted back due to the fact that the adversarial PS maintains a top priority in the PACS.

Passive injection systems such as the accumulators, cannot make up for the lack of availability of the rest of the systems. Any passively injected water will be eventually boiled and vented, ending up in the containment sumps through the burst rupture disks of the Pressurizer Relief Tank, but without entering into the recirculation mechanism, so it would not contribute to recover the coolant inventory.

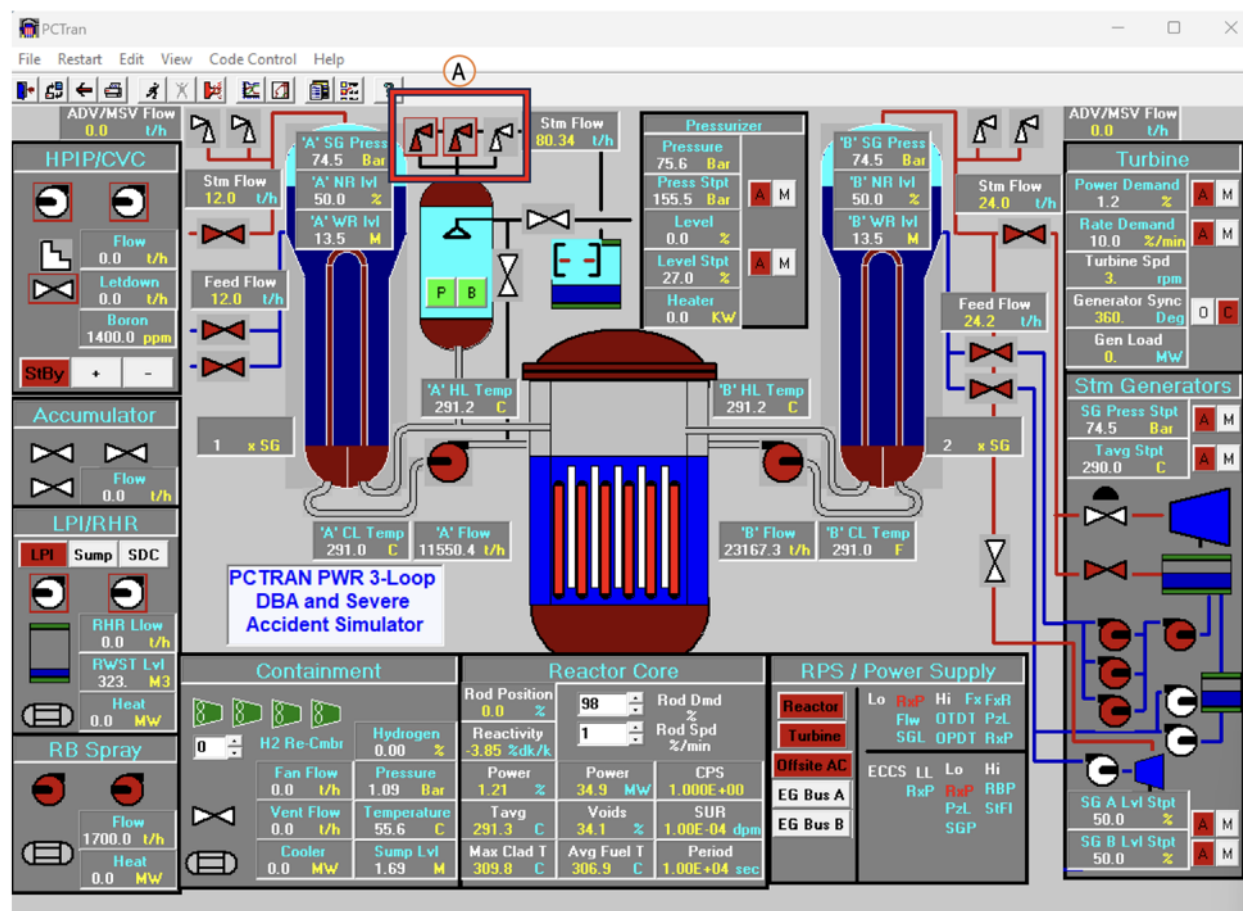


This situation poses a challenge to keep the core cooled and covered, as the SLOCA keeps progressing. Let's finish the analysis by simulating the cyber-physical attack to understand its potential consequences.

4.4.7 - Simulating the cyber-physical attack

This cyber-physical attack was simulated by using PCTran, a commercial tool used⁹⁰ to simulate severe accidents and Design Basis Accidents in nuclear power plants. I would like to clarify that I found this version (a 3-loop PWR), which was licensed for an Asian government, leaked in a github repository.

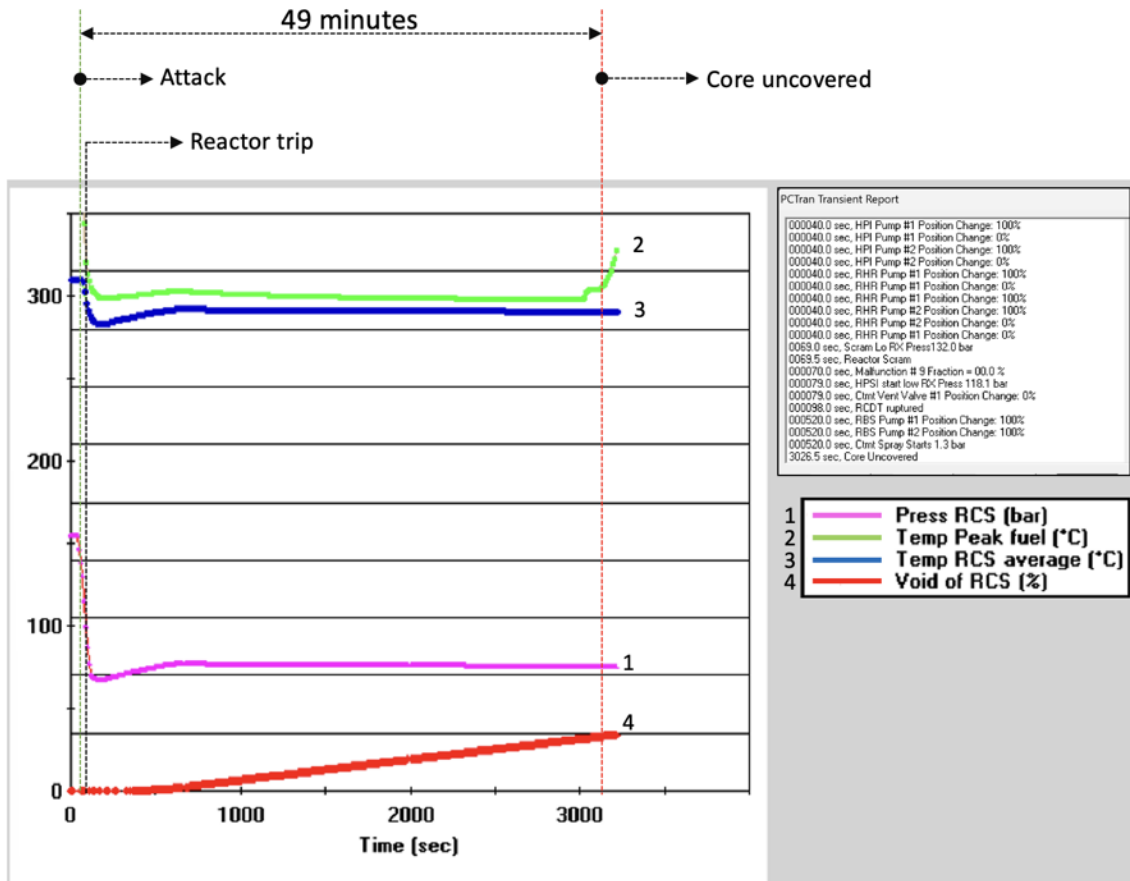
- A) (Beginning of Cycle) Two PSRVs were stuck open to simulate the initiating event in our cyber-physical attack. At the same time, the SIS and RHRS were inhibited



The following plot shows the progression of the cyber-physical attack, from the initiating event (2 stuck-open PRSVs) to the moment the core is uncovered, just 49 minutes later.

Let's analyze the sequence of events.

⁹⁰ <https://www-pub.iaea.org/MTCD/Publications/PDF/TCS-68web.pdf>



- 40s - Cyber-physical attack starts

The pressure in the RCS (1 – Press RCS) suffers a pronounced decrease due to the initiating event and the reactor shutdown. As the attack progresses the RCS pressure will reach an equilibrium with the pressure of the SGs, as it is the only working heat exchange mechanism (in addition to the 2 stuck-open PSRVs).

- + 69s - Reactor trip

Due to the low RCS pressure, the reactor is tripped. This significantly reduces the temperature both in the coolant and in the fuel. The reactor trip decreases the temperature of the Hot leg (as the chain reaction in the core has been stopped), which is where the pressurizer is connected, through the surge line. As temperature decreases in the Hot leg, the water level in the pressurizer falls due to water contraction. This makes the pressure fall even further.

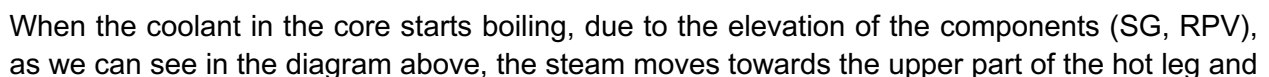
- + 79s - HPSI is inhibited

10s after the reactor trip, pressure in the RCS is now low enough to initiate the High-Pressure Safety Injection. However, as we have previously shown, this actuation is inhibited as a result of the cyber-physical attack.

The rupture disks of the Reactor Coolant Drain Tank (similar to the PRT) burst due to the continuous flow coming from the 2 stuck-open PSRVs. Now the coolant is being spilled out onto the containment floor, where the sumps are located. This increases the pressure inside the reactor containment.

Due to the increasing pressure inside the reactor containment, the Reactor Building Sprays (RBS) are activated to condensate the steam and reduce the pressure.

A limited amount (to prevent a positive Moderator Temperature Coefficient) of soluble Boron (Boric Acid), a neutron poison, is routinely added (via the CVCS) to the water to help cope with excess reactivity, especially during the beginning of the cycle (BOC), when new fuel has been loaded.



SG. As this steam is ascending into the U-Tubes, it will eventually condense, thus removing the soluble boron via distillation. This process will then begin accumulating deborated water in the lower parts of the SG and cold leg (pump seal).

Safety Injection systems inject borated water into the core, but as these systems have been inhibited, we are progressively creating a boron dilution scenario. This means that fewer neutrons are being absorbed thus adding positive reactivity. This is especially problematic during BOC, because the excess reactivity is expected to be controlled not only with the control rods but with the added boron. So even in shutdown mode, the negative reactivity added by the control rods would not be enough to offset the positive reactivity added by a boron-diluted moderator. This may create a reactivity-initiated accident which could damage the fuel.

However, back at the simulation plot, there is a much more pressing issue: the void fraction (4 - Void of RCS) in the coolant is dangerously increasing.

+ 3029s - Core uncovered

49 minutes after initiating the cyber-physical attack, the upper part of the core is uncovered. This means there is not enough coolant available to keep the fuel cool, so its temperature (2 - Temperature peak Fuel) initiates a brutal excursion.

Finally, the core is (partially) melting.

A core melt does not mean a guaranteed catastrophe, but it is a pretty serious scenario. In addition to persisting in trying to flood the core with water (borated or not, because with a melting core its geometry is now useless, so criticality is not the most pressing issue at this point), now almost everything depends on the containment.

As long as the structural integrity of the containment is maintained, the consequences of a core melt can be greatly mitigated. There are different structures, materials, and both passive and active systems designed for such a critical task. For those readers interested in this issue, the 'U.S. EPR Severe Accident Evaluation Topical Report'⁹¹ is highly recommended.

However, I am deliberately excluding anything related to this last stage from the scope of this research.

⁹¹ <https://www.nrc.gov/docs/ML0631/ML063100157.pdf>

5. Conclusions

Modern nuclear facilities increasingly rely on sophisticated digital Instrumentation and Control (I&C) systems to monitor, control, and safeguard reactor operations. This reliance on digital systems, while offering efficiency and automation, may introduce new attack vectors that can be exploited by malicious actors.

Despite international legal frameworks generally prohibiting attacks on nuclear facilities, the allure of targeting these critical infrastructure assets remains a concern in the context of nation-state conflicts. The potential for causing, through cyber means, widespread societal disruption, inflicting economic or military damage, and instilling fear makes nuclear facilities attractive targets for state-sponsored actors seeking to advance strategic objectives during armed conflicts or profound geopolitical confrontations.

In general terms Nuclear Power Plants, and especially Pressurized Water Reactors (PWRs), are designed with robust, multi-layered safety measures to prevent accidents and mitigate potential risks. A key principle underlying these designs is "Defense in Depth" (DiD), which incorporates multiple, independent layers of protection, including physical barriers, redundant and diverse safety systems, and strict operational procedures. This means that even if one safety layer fails, other layers remain in place to prevent a major incident.

As a result, the difficult task of attacking a nuclear reactor through cyber means requires, among other things, systematically analyzing potential weaknesses in digital Instrumentation and Control (I&C) systems. This research explored the potential for cyber-physical attacks targeting nuclear reactors, with a focus on Pressurized Water Reactors (PWRs).

The analysis centered on the security challenges presented by digital I&C systems, particularly the Teleperm XS (TXS) platform, widely used for implementing safety-critical functions like the Reactor Protection System (RPS) and the Engineered Safety Features Actuation System (ESFAS).

The ultimate goal was to identify and evaluate a sequence of physical events, triggered through digital manipulation, that can overcome the plant's defenses and potentially lead to a core melt, structural damages, or release of radioactive material.

While this research focused on potential cyber-physical threats to traditional nuclear power plants, it is important to consider the future of nuclear energy, particularly the development of Small Modular Reactors (SMRs). SMRs, with their enhanced safety features, reduced capital costs, and potential for deployment in remote locations, represent a promising path towards a sustainable energy future. However, as SMRs are expected to be increasingly integrated as critical infrastructures, it becomes paramount to prioritize cybersecurity considerations. The insights gained from analyzing potential issues in existing digital I&C platforms and NPP designs, may serve as valuable lessons for designing and implementing robust and resilient safety systems for SMRs, ensuring the peaceful and secure utilization of this vital energy source.

This research should be interpreted primarily as an informative endeavor rather than an alarmist one. Transparency and robust educational initiatives can help dispel myths and foster a greater public understanding of nuclear technology.

In a world facing increasingly sophisticated cyber threats, a more informed community, equipped with a nuanced understanding of both the risks and the robust safety measures in place, will be better prepared to deal with potential nuclear-related incidents.

About the author



Ruben Santamarta is a European independent security researcher with over 20 years of experience in the industry.

He has found and published dozens of vulnerabilities in a variety of targets, such as: desktop software and mobile apps, e-voting platforms, operating systems, Industrial Control Systems, SCADA software, IoT devices, RF controllers, satellite terminals, maritime equipment, avionics, and radiation monitoring systems.

His main areas of expertise are reverse engineering, source code analysis, and Industrial Control Systems.

Ruben has presented multiple times at international security conferences, such as Black Hat USA.

The best way to reach out is via LinkedIn. Please send a connection request, outlining the purpose of it in the message.

<https://www.linkedin.com/in/rubensantamarta/>



<https://www.reversemode.com>

This paper is released under the following license⁹²:



October 1st, 2024.

⁹² <https://creativecommons.org/licenses/by-nc/4.0/>

Appendix A

Figure 1 - Teleperm XS SCP3



Figure 2 - Teleperm XS SVE2

