ISSUE: February 2020

Transmitting Overpower Alert Signals With Low Latency Boosts Reliability Of Bus-Bar Power Lines

by Viktor Vogman, Power Conversion Consulting, Olympia, Wash.

Server power supplies are typically sized larger than strictly necessary in order to deliver more power than a system may initially need. Most server systems are populated with fewer hardware components than their projected capacity. Further, in the vast majority of cases, not all the components used are the most power hungry.

For example, many systems use slower processors that usually consume less power than the fastest processor in their class. On top of this, the CPU utilization reaches its 100% capacity relatively infrequently—and when this does happen, this mode of operation can be effectively supported by the peak power ratings of the power supplies. On average, many systems consume power levels far lower than the projected peak, and hence could function adequately and more efficiently with a smaller, lower-rated power supply.[1,2]

The above discussion concerns server systems containing one or several motherboards each and having a dedicated ac-dc power supply, or possibly two ac-dc power supplies for redundancy. Such a server rack receives ac power.

However, for server racks with dc bus-bar power distribution, where one common power subsystem supplies dc power to many server nodes, the power supply is usually sized for a full-rack configuration running software that generates the highest power consumption. This results in a larger continuous wattage rating for the power supply and a higher cost for the server rack. Such overly conservative power delivery architectures present an opportunity for significant reduction in power supply size and cost and for improvement of system performance-per-watt scores.

When operating in a failure/nonredundant mode, smaller power supplies can detect an anomalous excessive power condition. They can then generate a fast interrupt in the form of an overpower alert that signals to the servers to throttle back until the redundancy gets restored and/or power comes back into an acceptable range.

Existing power throttling solutions employ separate, daisy chain wiring to carrying such interrupt signals. But these wire assemblies are expensive and unreliable due to lack of redundancy. For example, opening any connection in the daisy chain cable may result in loss of key rack features such as overpower protection, power capping, and the ability to successfully ride through ac line voltage dropouts. Opening a connection in the signal daisy chain may also decrease system availability due to the possibility of server nodes crashing when the power subsystem is overloaded.

Thus, providing fast load control in the power subsystem without using unreliable separate daisy chain wiring represents an attractive option for protecting a shared power subsystem and the rack from crashing. Developing a technique for sending an alert logic signal from rack power supplies to the multiple rack nodes receiving dc power over the same power delivery bus plays a key role in implementing such a protection method.

A number of IC manufacturers have developed transceiver ICs for power line communication (PLC). These ICs can be effectively used for power management needs, such as balancing power in the rack servers, power capping, or controlling each node in a rack so that it consumes only an allocated portion of the delivered power. However, latencies in these PLC systems associated with signal processing prevent their use for fast power throttling and power subsystem protection.

This article presents a study of the conditions necessary for reliable transmission of overpower alert signals across a dc bus-bar power delivery network (PDN) in a high-power server rack. It also introduces a simple technique for broadcasting the alert logic signal across such PDNs with a small-size transmitter/receiver and with minimal latency in node throttling.

Factors Affecting Alert Signal Throughput

There are several factors impacting communication between shared rack power supplies and multiple server nodes. Both PSU output and node power input typically have large decoupling capacitors, which present very low impedances for high frequencies or for high-slew-rate signals. Generating meaningful signal levels across such components would cause the supplied voltage disturbances and require high-power transmission stages, which would offset the advantage of transmitting interrupt signals over the power lines.

Typical PLC systems also impose a requirement of providing a specified minimum power line impedance at the PLC carrier frequency. This is needed to ensure that the transmitter will not be overloaded and that an appropriate signal level will reach the receiver input. To prevent low impedances from impacting the transmitter power rating, a dc power distribution arrangement needs to incorporate buffer elements on both ends of the power line.

Inductors are the most commonly used components for such buffers. These components are not only capable of increasing the impedance of the line, but also can greatly benefit power line protection by causing the faulty node current to ramp up in a fully controllable manner. The inductor size must be large enough to provide sufficient impedance for the conventional power-line-communication signal, such that the power supply and the nodes do not create low-impedance paths for this PLC signal and do not interfere with it.

Protection of the common bus-bar PDN from node faults is provided with hot swap controllers (HSCs) detecting severe overcurrent, which is an indicator of a short-circuit condition. A fast response circuit in an HSC can usually control the current within 1 µs after the threshold is crossed. But in absence of buffer inductors or if a buffer inductor value is not adequate, even such a fast response may not prevent excessive bus current and an out-of-spec condition for the bus voltage. This is illustrated in Fig. 1.

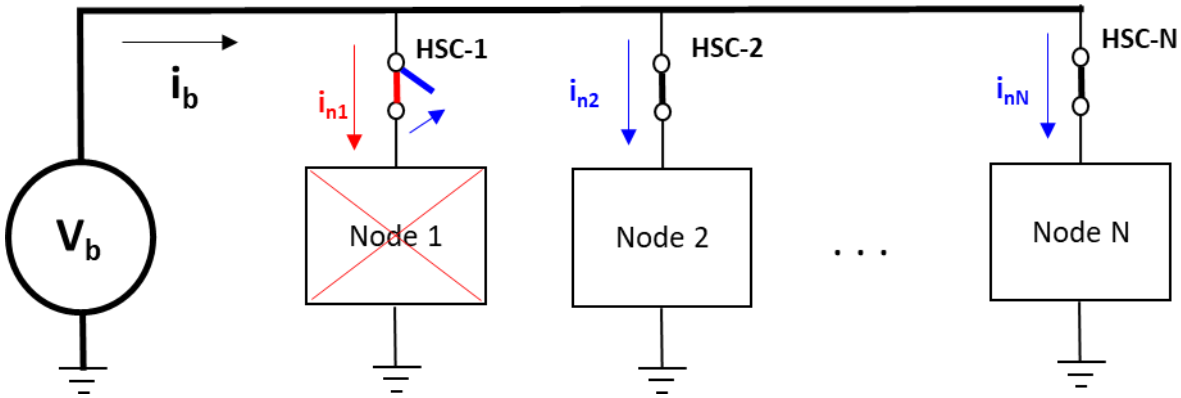


Fig. 1. Simplified block-diagram of a common bus-bar dc PDN supplying power to N nodes. Protection of the PDN from node faults is provided with hot swap controllers detecting severe overcurrent and isolating the faulty load from the PDN. In absence of buffer inductors the current in the faulty node (i_{n1}) and bus current i_b can reach excessive levels and even a very fast HSC response may not prevent excessive bus current and an out-of-spec condition for the bus voltage V_b.

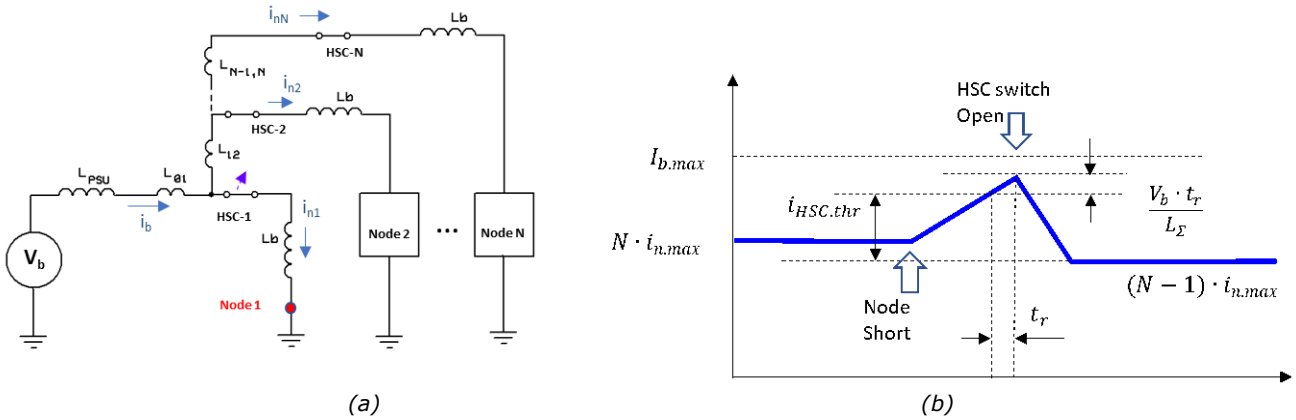
In other words, in the absence of a specified series inductance in a short circuit path the node fault impact on the bus appears to be undetermined. This means that the inductor can also play a key role in buffering a faulty node until the HSC properly isolates it from the common bus. Thus, the problem of developing an alert signal broadcasting technique can be split into two sub-problems. This first is defining requirements for the buffer elements that manage node-fault conditions. The second is selecting actual transmitter and receiver hardware that operate reliably with the specified buffers.

Selecting Buffer Inductor Values For Node Fault Management

As stated above, the node-side and power-supply-side buffer inductors are required to provide a specified minimum impedance at the PLC frequency. These inductors also play a key role in node fault management because they limit the ramp rate of the faulty load current and can make the power line behavior under a load fault condition fully predictable.

A simplified equivalent circuit of a PDN delivering dc power to *N* nodes and having a node fault is shown in Fig. 2a. In this circuit L_{PSU} and L_b represent power supply and node buffer inductors, respectively, and L_{01}, L_{12}, etc. are the stray inductances of different bus-bar sections. The HSCs are represented by closed switches HSC-1, HSC-2, etc.

Let's consider the worst-case scenario where all but one node operate at their maximum currents $i_{n.max}$ and node 1, closest to the power subsystem, fails into short circuit (Fig. 2a). Its hot swap controller switch HSC-1 will disconnect it from the PDN a short time t_r after the HSC overcurrent threshold $i_{HSC.thr}$ is crossed (Fig. 2b).



(a) (b)

Fig. 2. Simplified equivalent circuit of a PDN delivering dc power to N nodes with node 1 failed into short circuit (a) and a PDN current diagram after a short circuit occurs in one of the nodes (b). The PDN current ramps up after a short and ramps down once the affected HSC switch isolates the faulty node from the bus.

In normal operation of a PDN the absolute max rated bus (power subsystem) current $I_{b.max}$ must not be exceeded. Thus, the sum of all load currents under a fault condition in one node, is described by the inequality:

$$(N-1)\cdot i_{n.max} + i_{HSC.thr} + \frac{1}{L_\Sigma}\int_0^{t_r} v_b(t)dt \le I_{b.max}, \qquad (1)$$

where $i_{n.max}$ is the max nominal node current, L_Σ is a total series inductance in the fault path (in our case $L_\Sigma = L_{PSU} + L_{01} + L_b$), and $v_b(t)$ is instantaneous bus voltage. Bus bar inductance values, expressed via per-unit-length inductances, are represented by the variable L_Σ and required for computing stray inductances L_{01}, L_{12}, etc. For vertically spaced parallel bus bars, or other transmission line geometries, L_Σ can be computed using the online inductance calculator.[3]

The inequality in equation (1) enables us to determine a requirement for the total series inductance L_Σ needed for reliable load fault management. Let's assume that, for a PDN load not exceeding the power line max rating, any change in bus voltage will be negligible compared to the nominal value of the bus voltage. In other words, if $v_b(t) = V_b = \text{const}$, we find that in order to meet inequality (1) L_Σ must satisfy the following condition:

$$L_\Sigma \ge \frac{V_b \cdot t_r}{I_{b.max} - (N-1)\cdot i_{n.max} - i_{HSC.thr}} \qquad (2)$$

Then, to ensure reliable communication, any reduction in the PLC signal magnitude must be insignificant, and therefore the buffer inductance values must be significantly larger than the cumulative bus bar stray inductance in the signal flow path. If the L_{PSU} and equivalent node buffer inductance (L_b/N) values are selected to be equal, such that $L_{PSU} = L_b/N$, then the minimum buffer inductance value that provides node isolation from the bus before the bus current crosses $I_{b.max}$ level can be determined from (2) as follows:

$$L_{b.min} = \frac{V_b \cdot t_r \cdot N}{[I_{b.max} - (N-1) \cdot i_{n.max} - i_{HSC.thr}] \cdot (N+1)} \qquad (3)$$

Equation 3 indicates that the required buffer inductance value is directly proportional to the response time t_r of the HSC IC. The shorter the t_r, i.e. the faster the hot swap controller, the smaller the buffer inductor required.

Note that equation 3 determines the minimum buffer inductance value that guarantees a safe fault isolation only. The actual buffer inductance value L_b needs to be selected as the larger of two values—one value is the result of equation 3 and the other is the inductance that provides the minimum specified power line impedance at the PLC carrier frequency. Having defined the requirements for buffer inductance values required for reliable node fault management we can now examine how the buffer inductance will affect transmission of the alert signal.

Considerations For Alert Signal Transmission

A functional block diagram illustrating alert signal transmission across the dc power line or rack dc bus bar PDN is shown in Fig. 3.

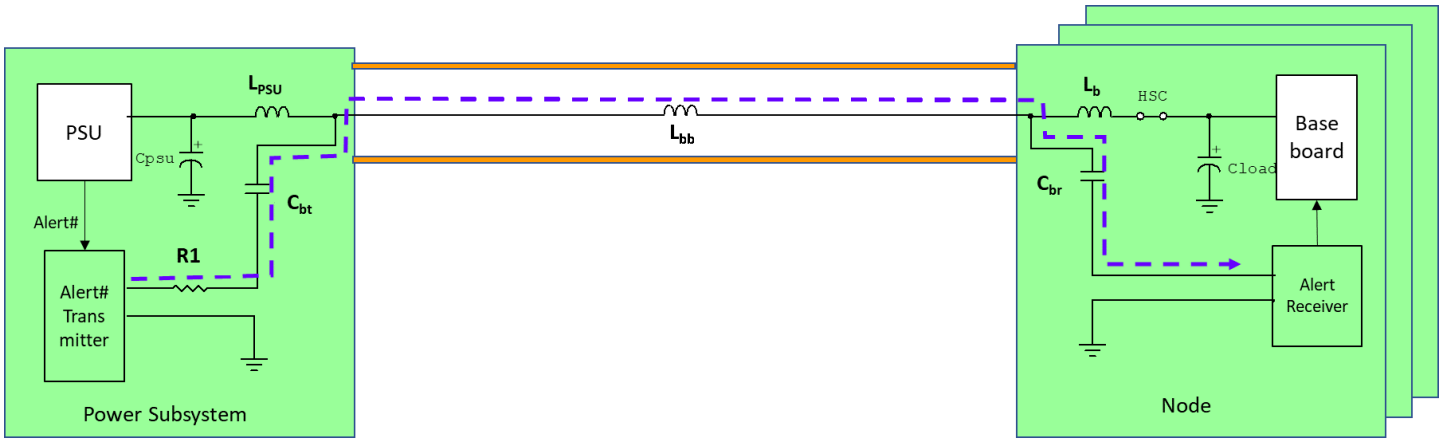


Fig. 3. Functional block diagram illustrating transmission of the alert signal across the dc power line.

In this diagram the transmitter operating on the PSU side and the receiver acquiring the alert signal on the node side are connected to the power line, which typically represents a low bus voltage (V_b < 60 V) PDN. The alert signal propagation path is represented by the blue dashed line. Similar to Fig. 2, the hot swap controller IC is represented by a closed switch labeled HSC.

In this diagram both the transmitter and the receiver are coupled to the PDN without galvanic isolation, through dc blocking caps C_{bt} and C_{br}, respectively. C_{PSU} and C_{load} represent decoupling caps on the power subsystem and the node sides, respectively. Inductance L_{bb} represents the cumulative bus-bar stray inductance between the PSU and the node.

Provided that the receiver decoupling cap C_{br} is large enough that its voltage change is negligible within the time interval of the alert pulse and that the receiver input impedance is much greater than the impedance of

the buffer element, we can construct the equivalent linear RLC circuit model by representing the PDN as a transmission line carrying the alert signal applied to its input. This is shown in Fig. 4.

This circuit model can be used to analyze the alert signal propagation through a power delivery network as through a linear circuit with R-L-C components and with output v_r taken across the equivalent buffer inductor $L_{b.eqv} = L_b/N$ at the receiver end (Fig. 4). If transmitter galvanic insulation from the bus is provided with a pulse transformer its leakage inductance needs to be added in series with the transmitter internal resistance represented by resistor R1.

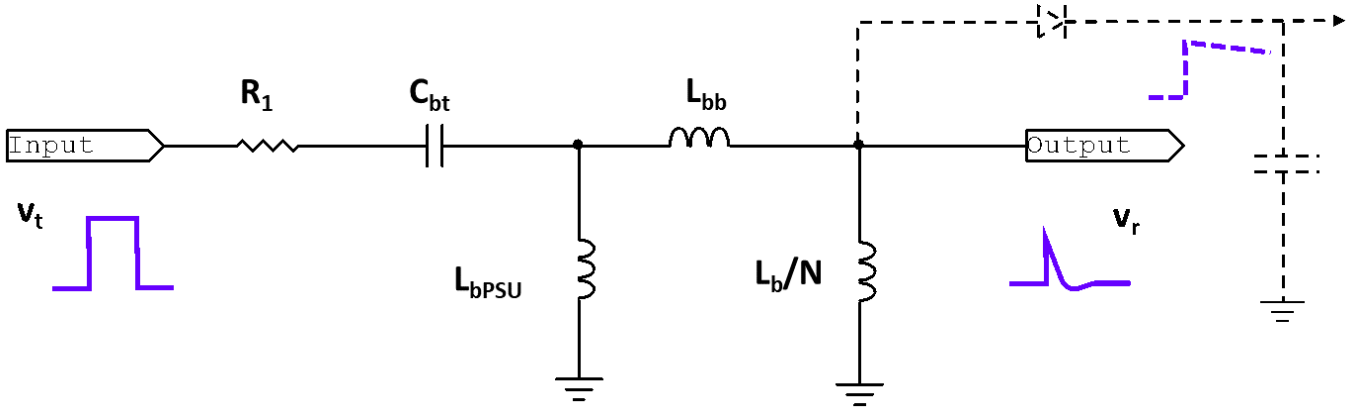


Fig. 4. An equivalent linear RLC circuit model can be used for analysis of transmission line voltage and current waveforms. The pulse duration at the receiver end can be extended with a peak detector having a needed hold time (dashed lines). For a version of the transmitter with transformer output, transformer leakage inductance needs to be added in series with R1.

A general case expression for the receiver voltage in the Laplace domain can be derived from its transfer function, yielding the following result:

$$v_r(s) = \frac{v_t(s) \cdot \frac{L_b}{N} \cdot s \cdot L_{PSU} \cdot (L_{bb} + \frac{L_b}{N}) / (L_{PSU} + L_{bb} + \frac{L_b}{N})}{\left[R_1 + 1/sC_{bt} + s \cdot \frac{L_{PSU} \cdot (L_{bb} + \frac{L_b}{N})}{L_{PSU} + L_{bb} + \frac{L_b}{N}} \right] \cdot (L_{bb} + \frac{L_b}{N})}$$

There is one undefined variable in this equation—the transmitter output (transmission line input) signal, $v_t(s)$. For the general case, this signal waveshape is unknown, so there is no universal analytical solution for v_r in the time domain. The equations describing instantaneous voltages and currents in this circuit for the v_t signal represented by the unit step function V_t/s (where V_t is the transmitter pulse magnitude) are widely used in basic linear electric circuit theory and will not be given here due to their complexity and difficulty of analysis.

However, the equivalent circuit model in Fig. 4 or its transformer-output version can be used with general Spice-based simulators with the input signal represented by a piecewise-defined function. An example of such usage with a common trapezoidal-shaped input signal is given in Fig. 5.

The simulation in Fig. 5 shows that if the output voltage short pulse (yellow trace, receiver input voltage, vr(t)) can be detected by the receiver there is no need for generating and transmitting a long-duration pulse via the power line: generating such a pulse would require a larger size buffer inductor L_b and larger transmitter stage.

It is more expedient to transmit the shortest possible duration pulse over the PDN and restore the needed alert duration at the receiver end. A fast comparator is usually capable of detecting a short (e.g. submicrosecond) voltage pulse at the receiver input.

The other option is to use a peak detector with sufficient hold time. This would "expand" the pulse, with some small delay, lengthening duration at the receiver-end, which allows the receiver to be configured with a general-purpose comparator instead of a fast comparator. This detector is represented in Fig. 4 with dashed lines.

The peak detector signal option is illustrated by simulation of the processes in the transmission line with the following parameters: L_{bb} = 400 nH, $L_{PSU} = L_b$ = 1 µH, N = 3, C_{bt} = 0.2 µF and R_1=3 Ω, with simulation results shown in Fig. 5. The green waveform shows the peak detector output, $V_{PD}(t)$, "expanding" the original pulse (the yellow waveform, $v_r(t)$) at the receiver end.

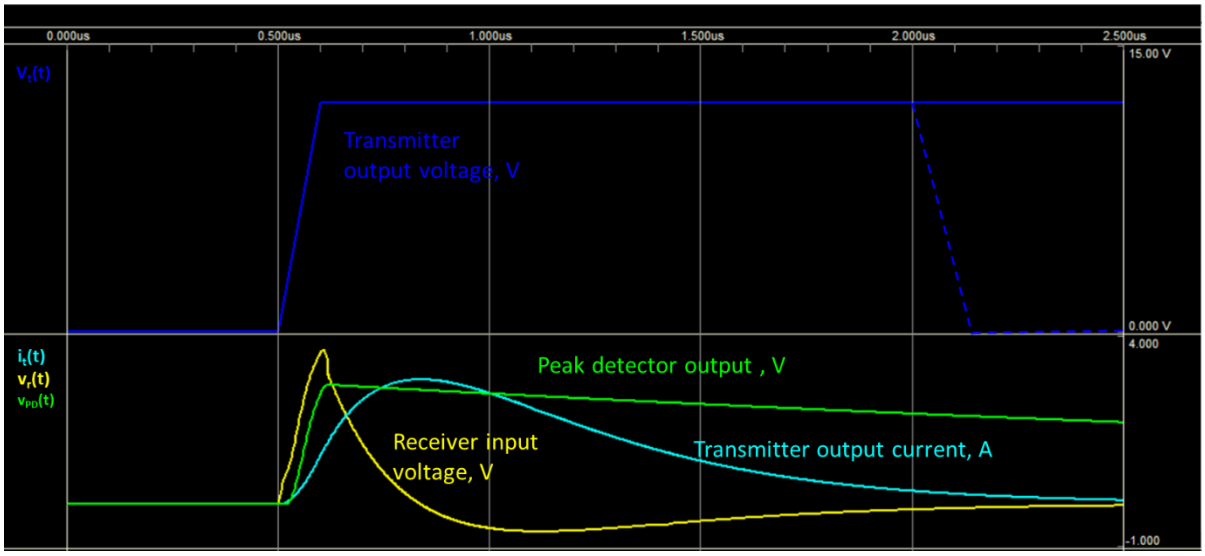


Fig. 5. Simulation of the transmission line modeled in Fig. 4 (with L_{bb} = 400 nH, $L_{PSU} = L_b$ = 1 µH, N = 3, C_{bt} =0.2 µF, v_t rise time = 100 ns and V_t = 12 V). An overdamped transmission line response provides a single short-duration voltage pulse at the receiver side (yellow trace). If the comparator response is not fast enough to detect it, the pulse duration can be extended with a peak detector having a needed hold time (green trace). Transmitting the shortest possible duration pulse over the PDN (dashed blue line) does not impact the alert delay but helps to reduce the transmitter size.

A desired pulse duration at the receiver end can be facilitated by a one-shot monostable multivibrator. Transmitting a single short pulse and increasing the value of R1 allows us to minimize the size of the alert signal transmitter and combine it with the PLC stage in one IC, if necessary.

Imposing assumptions on the transmitter-generated signal waveshape, as applied to the circuit in Fig. 4, allows us to conduct a "ballpark" feasibility analysis for typical cases and to offer some practical recommendations.

The value of passive resistor R1 in this circuit determines whether the transient process will have an oscillatory behavior associated with generating multiple voltage pulses at the receiver input. To prevent this oscillation from happening, select R_1 such that it provides an overdamped transient response, producing a single pulse at the receiver end, as shown in Fig. 5. Thus, R_1 should meet the following criterion:

$$R_1 > 2\sqrt{\frac{L_{PSU}//(L_{bb} + L_b/N)}{C_{bt}}}$$

With an overdamped condition the transmission line response is very similar to the R-L differential circuit response, which makes the duration of the pulse generated at the receiver input shorter than the duration of the transmitter pulse.

As it can be seen from the circuit topology in Fig. 4, series components R1 and C_{bt} can significantly impact the signal magnitude V_r at the receiver input, making it too small for reliable detection. This impact will be especially noticeable for slow transmitter-signal ramps. To prevent too small a signal for V_r, the leading edge of the transmission pulse should be made as short as possible.

With fast transmitter voltage v_t ramps ($t << L_{PSU}/R_1$, $t << R_1C_{bt}$) the transmission line during the short ramp time interval essentially represents a voltage divider formed by L_{bb} and L_b/N. In this case the signal magnitude at the receiver input V_r can be determined with good (for practical purposes) accuracy as:

$$V_r = V_t \cdot \frac{L_b}{N \cdot \left(L_{bb} + \frac{L_b}{N}\right)}$$

where V_t is the transmitter pulse magnitude. This signal magnitude needs to be 1.5 to 2.0 times greater than the PLC signal magnitude (typically 1 V):

$$V_r \geq (1.5 \ldots 2) V_{PLC}$$

Meeting this condition helps the alert receiver easily distinguish between these two signals and detect the alert. Galvanically isolated transmitter and receiver implementation examples are given in reference [4].

A Proof-Of-Concept Test Setup And Experimental Results

A proof-of-concept setup was constructed that was basically equivalent to the block diagram in Fig. 3. The only difference in the setup was the galvanic isolation provided by the Murata 5300 series miniature isolating pulse (current) transformers at the transmitter output and receiver input. Major components used in this setup and their descriptions are given in the table below.

Table. Components and specifications for PDN prototype used to measure performance of alert signal method.

Component / Parameter	Description
Power subsystem	12 Vdc, 1.1 kW; 1.4 kW peak
PSU buffer inductor	L_{PSU} = 200 nH
Bus bars	Two parallel bus bars (W x T x L = 3 in. x 0.12 in. x 38 in.)
Bus bar center-to-center spacing	D = 0.75 in.
Number of e-load emulating nodes (N)	3
Node buffer inductors (selected using Eq.2)	L_{b1} = L_{b2} = L_{b3} = 1 µH
HSC	VT505FQX protection IC with an integrated MOSFET
Transmitter output transformer	53100C
Receiver input transformer	53040C

The experiment was conducted in two stages:

1. Verification that the bus-bar power distribution line can be reliably protected with the selected buffer inductor and HSC IC.

2. Stressing a power line with a slammer transient tool[5] to generate an alert signal and measuring the delay between the leading edges of the transmitter and receiver alert signals.

Experimental waveforms from the power line protection process are shown in Fig. 6. The HSC output was crowbarred with a 10-mΩ solid-state switch to emulate a fault (node short circuit) condition. The HSC switch opened in about 8 µs after the short was applied.

With the selected 1-µH buffer inductor the bus current ramped up in a controlled manner (practically linearly with $di/dt = V_b/L_b$). At the time when the switch isolated the faulty load from the power line, the short circuit current did not exceed 105 A, while the power line voltage remained within its spec limits.

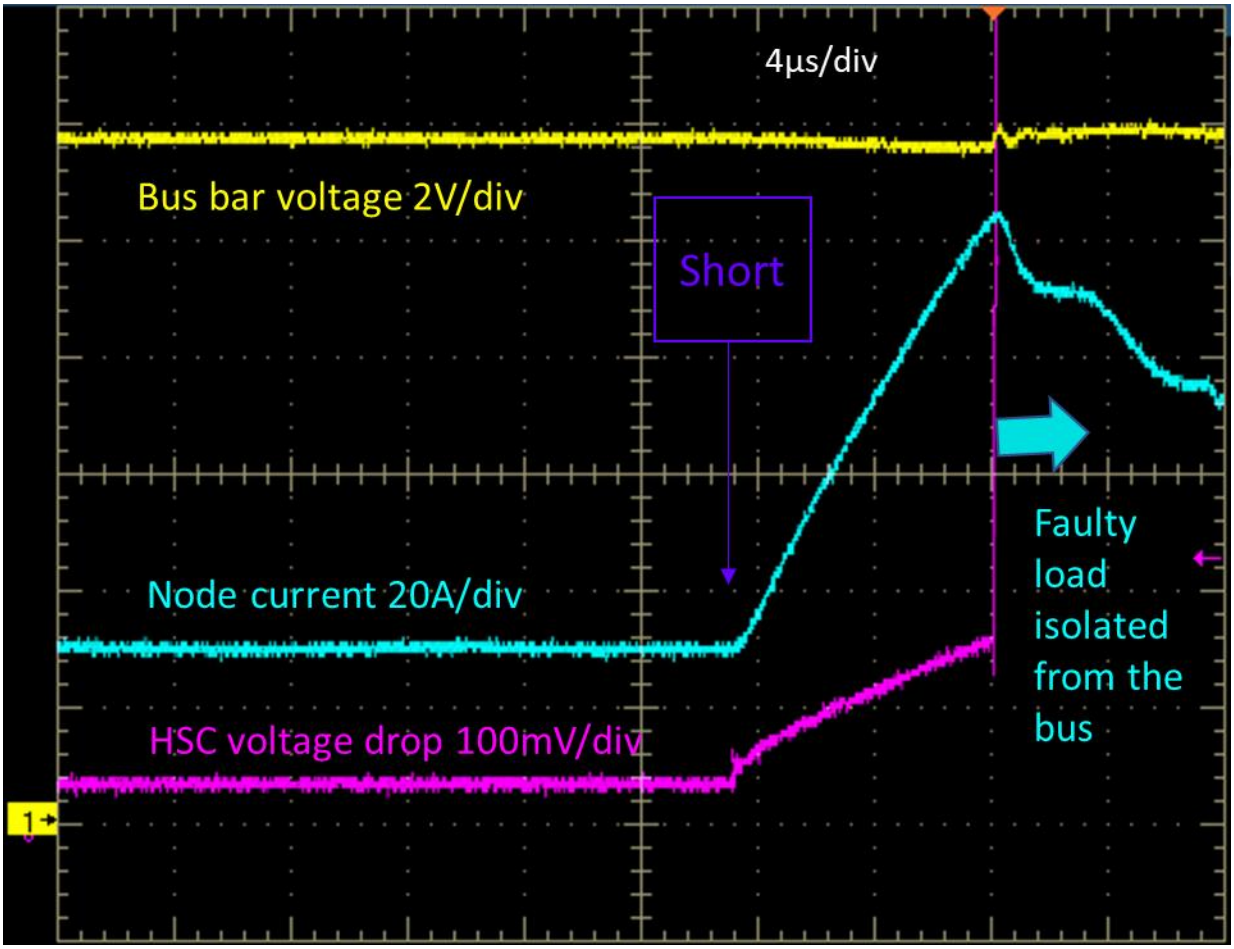


Fig. 6. With the selected 1-µH buffer inductor, the current ramped up in a controlled manner and at the time the HSC switch started isolating the faulty load from the power line, the short circuit current reached its maximum (105 A) while the power line voltage remained within its spec limits.

Experimental measurements of the alert signal latency are given in Fig. 7. The PSU was stressed with a slammer transient tool, which applied a current pulse in addition to the current drawn by the nodes. In response the PSU generated an alert signal (green trace in Fig. 7) warning that the peak PSU rating had been reached and load power needed to be reduced. The current waveform of the alert signal indicated that the

response was overdamped and the three receivers' alert signals changed their states within a 10-µs time interval.

This delay is considered acceptable for most power supplies, as it can be supported by their output decoupling caps. If the transmitter and receiver are coupled directly to the bus bars then transformer leakage inductance would not be impacting signal propagation and the delay can be significantly reduced.

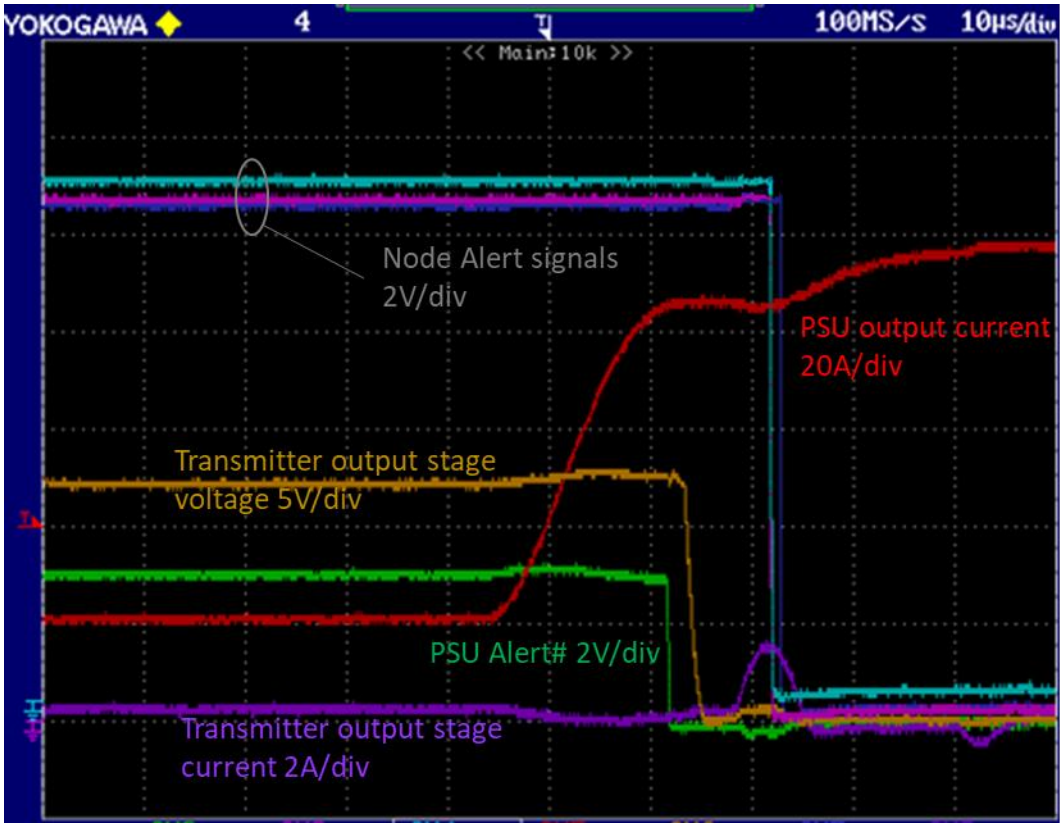


Fig. 7. In the overdamped mode only a single short current pulse is generated in the transmission line, which allows it to reliably generate alert signals in multiple receivers within a short (10-µs) time interval.

Conclusions

The proposed alert signal broadcasting technique can be used in combination with existing PLC systems or integrated into the PLC ICs for their enhancement. To do so, the designer must observe the following requirements.

Buffer inductors required for power line communication, positioned on the PSU and node sides, serve two additional purposes. One is to keep the power line current below the rated level under a node short-circuit condition (safe node fault isolation). The other is to provide the required high impedance for broadcasting of logic signals across the line. The buffer inductance value needs to be selected as the greater of the values that guarantee safe node-fault isolation and that provide the specified power line impedance at the PLC carrier frequency.

An equivalent linear circuit model of the power line can be used to select transmission line parameters and to analyze the line signal (voltage and current) waveforms. The shortest alert propagation delay can be achieved with direct coupling of the transmitter and receiver to the power line. Selecting transmission line components

that provide an overdamped transient response prevents the generation of multiple pulses at the receiver end. Using a peak detector with sufficient hold time helps to ease requirements for the receiver alert detector.

References

1. Optimizing Transfer Switch N+1 Redundant Power Architectures" by Viktor Vogman, How2Power Today, May 2019.
2. "Energy-Based Efficiency Metric Helps To Optimize Server Power Delivery For Dynamic Workloads" by Viktor Vogman, How2Power Today, October 2018.
3. Inductance Calculator.
4. Methods and systems for server power line communication, US Patent 9461709B2, 2016.
5. "Using Mini E-Loads To Emulate Fast Slew Rates And High Peak Power Levels" by Viktor Vogman, How2Power Today, September 2019.

About The Author

Viktor Vogman currently works at Power Conversion Consulting as an analog design engineer, specializing in the design of various power test tools for ac and dc power delivery applications. Prior to this, he spent over 20 years at Intel, focused on hardware engineering and power delivery architectures. Viktor obtained an MS degree in Radio Communication, Television and Multimedia Technology and a PhD in Power Electronics from the Saint Petersburg University of Telecommunications, Russia. Vogman holds over 50 U.S. and foreign patents and has authored over 20 articles on various aspects of power delivery and analog design.

For more information relating to powering servers, see How2Power's Design Guide, locate the "Application" category and select "Data Centers".